Essays on Information and Collective Choice

by

Sukanta Bhattacharya

B.Sc., The University of Calcutta, 1992
M.Sc., The University of Calcutta, 1994

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

July 2006

Abstract

This thesis consists of three essays that attempt to contribute towards a better understanding of the collective choice problems in presence of incomplete information.

The first essay of this dissertation investigates a voting model with conflict of interest between the designer of the voting rule and the voters who share a common preference. We show that, given sophisticated voting by the voters, designer's optimal voting rule either nearly coincides with the voters' optimal rule or is a near unanimity rule for one of the alternatives. When the designer has a very strong bias in favour of one of the alternatives, her best option may lie in increasing the threshold of votes for that alternative.

The second essay discusses an indirect voting mechanism that can achieve better informational efficiency. It is well known that under usual (simultaneous) voting rules, private information held by voters are imperfectly aggregated. We consider a multi-round sequential voting procedure which allows voters to *choose* when to cast their votes. Without any conflict of interest among voters, there is always an equilibrium under this rule which perfectly aggregates all the available information. Moreover, in an environment with conflicting interest among voters, we show that this indirect mechanism achieves as an equilibrium outcome what pre-play communication can achieve.

The third essay examines committee design under endogenous information and shows that two opposing effects - the *free-rider* effect and the *information complementarity* effect - could get intensified in different parameter regions as the committee size increases. This induces a trade-off between quality and quantity of information for the committee designer. The model identifies parametric situations where it may be optimal to create a smaller committee to ensure better quality of information.

Contents

Abstract ii

Contents iii

List of Figures v

Acknowledgements vi

Dedication vii

1 Overview and Summary 1

2 Optimal Voting Rules with Divergent Preferences 9
- 2.1 Introduction 9
- 2.2 The Model and Results 17
- 2.3 Concluding Remarks 36

3 Information Aggregation in Multi-round Elections 38
- 3.1 Introduction 38
- 3.2 The Model with Common Interest 46
- 3.3 Diverse Preference 56
- 3.4 Concluding Remarks 68

4 Committee Design with Endogenous Information 70
- 4.1 Introduction 70
- 4.2 The Model 76
- 4.3 Committee Design 80
  - 4.3.1 One member committee 81
  - 4.3.2 Two member committee 82

4.3.3 Comparison between committees 87
4.4 Concluding Remarks 95

Bibliography 96

A Appendices 103
A.1 Appendix to Chapter 2 103
A.1.1 Proof of Lemma 2.3 103
A.1.2 Proof of Lemma 2.4 108
A.2 Appendix to Chapter 4 110
A.2.1 Proof of Lemma 4.3 110
A.2.2 Proof of Lemma 4.4 112
A.2.3 Proof of Proposition 4.1 112
A.2.4 Proof of Proposition 4.2 115

List of Figures

3.1 Equilibrium path of play for class 0 equilibria in DGS 60
3.2 Equilibrium path of play in multi-round voting game 67

4.1 Pure Strategy Equilibria for Different Prior Values: $c < c_1$. . . 90
4.2 Pure Strategy Equilibria for Different Prior Values: $c > c_2$. . . 91

Acknowledgements

I must express my indebtedness to my supervisor Professor Parikshit Ghosh without whose guidance and supervision I wouldn't be able to complete this thesis. I am also grateful to Professor Patrick Francois and Professor Okan Yilankaya who, as members of my supervisory committee, have always been sources of helpful suggestions and inspiration. I would also like to thank Professor Ashok Kotwal, Professor Mukesh Eswaran, Professor Yoram Halevy, Professor Gorkem Celik, Professor Francisco González and Professor Art Shneyerov for useful comments at various stages of the dissertation. I had many illuminating discussions with my colleagues at the University of Calcutta and my classmates here at the University of British Columbia which helped me in different, often indirect, ways in writing this thesis. Finally, I must acknowledge the contribution made towards this thesis by my family, especially my wife Piyali Sarkar, who had to bear with all the family responsibilities during my long absence away from home.

To My Wife Piyali

and

Daughter Sharanya

1. Overview and Summary

This dissertation consists of three essays that contribute towards the understanding of collective choice problems in presence of incomplete information. Any acceptable social choice rule should address two separate concerns. The first question deals with finding social choice rules that leads to an acceptable aggregation of individual preferences. The literature on welfare economics has dealt with this issue extensively. But any study that aims to deal with efficiency of collective decision in an uncertain environment must also acknowledge the role private information can play. Individual choice depends on private information and individual decisions are somehow aggregated to arrive at a collective decision. More informed decision-making increases the overall efficiency of the outcome. An efficient social choice rule, therefore, besides finding an acceptable way to aggregate (possibly) diverse preferences of individuals, should recognize the importance of private information and take into account two factors. Institutions concerned with collective decisions should be designed to give proper incentives for information collection to the agents. Secondly, the private information of the agents should be used in the best possible way while making the final decision. These are the issues this dissertation intends to address.

Voting is one of the most common social choice mechanisms used in practice to arrive at a collective decision. One purpose of holding elections is clearly to aggregate individual preferences. But even abstracting from that, differences of opinion among agents may arise due to different private information. Hence, one purpose that an election intends to serve is to aggregate this disparate information. It has been recognized in the literature that there are some voting mechanisms which serve this purpose better than others. Now, assuming that society's average preference is the same as that of the body to which the decision is delegated (e.g. a jury or a committee), the voting rule that leads to best possible use of information is optimal for the society. In the first essay of this dissertation we take up an abstract model of voting to examine the consequence of conflict of interest among the society at large and members of the jury or the committee. We characterize the optimal voting rule from society's point of view.

The second essay of the thesis deals with the issue of information aggregation more closely, again in the context of a voting model. A well-known result in the literature on strategic voting is that simultaneous voting procedures cannot lead to efficient information aggregation if the information structure is rich enough. We look for a richer, yet practical, voting procedure that can achieve this end. We show that if the voters are allowed to choose the timing of their votes, then, under any voting rule, a substantial improvement towards better information aggregation can be made.

The third and final essay of the thesis travels in another direction to investigate the incentive for collecting information when information is endogenous. We provide a model with multiple agents where agents are allowed to choose both the quality and the quantity of information. We show that there may arise situations where a smaller group performs better even under the assumption of full information aggregation.

The analytical framework adopted for this study has some restrictive features which are worth spelling out to give the reader a sense of its scope and possible limitations.

First, we restrict attention throughout to binary choices. While restrictive, many problems of collective choice do involve two choices. For example, whether a bill should be accepted in the legislature, a defendant should be convicted on criminal charges, a job applicant should be hired, a particular Presidential candidate should be chosen in a two-party system, whether a medical panel should recommend surgery for a patient, and so on.

Second, we adopt a preference structure where all conflicts of interest (if any) arise from the issue of how to deal with uncertainty. In other words, in our framework, if the state-of-the-world were commonly known, everyone would agree about the correct course of action. Since only noisy information is available, preferences can be characterized by the relative weights agents place on the two types of errors that may arise. Most of the analysis goes further and assumes that agents also share a common distribution of weights. Preferences of this class, while special, seem reasonable descriptions of some real world problems. For example, people agree that murderers should be imprisoned and innocent people are entitled to personal freedom. Many economic policies (e.g. trade) have distributional consequences and therefore lack of consensus regarding the optimal choice. However, disagreements over monetary policy seem to be rooted more in disagreements over potential consequences rather than conflicting objectives (everyone benefits from low inflation and reduced macroeconomic fluctuations). Members of academic departments have substantial common interest in hiring superior candidates and rejecting inferior ones, even though private information may lead them to rank the candidates for a position differently. In these scenarios, the problem is not how to resolve the tension between conflicting objectives, but how to aggregate private information and make collective choices as responsive to available information as possible.

Many interesting problems, on the other hand, clearly involve conflicting aims as well as dispersed information. For example, people's preferences over Presidential candidates are likely to be a function both of perceived ability and the voter's ideological leanings. It would be nice to develop a theoretical framework that allows for both these aspects and their interplay. However, this is not an easy analytical task and the literature based on such analysis is still nascent. We believe, along with many other researchers in this area, that the insights learnt from studying the *pure* information aggregation problem can be useful in studying the more complex, hybrid cases.

Chapter 2 analyzes the properties of collective decision making when it is implemented through some voting rule that gives each agent a single transferable vote and a minimum threshold of votes necessary for each alternative to be chosen. This is a very limited class of mechanisms. For the environments and problems considered, more general mechanisms will typically lead to more efficient outcomes. In the real world, overly complex mechanisms may be too costly or place too high a cognitive demand on participants, possibly outweighing their benefits in many cases. In any case, voting is a commonly observed method of making collective decisions, and therefore understanding its strategic and welfare consequences is of practical relevance.

Since voting is often an imperfect way of aggregating information, there is also value to asking normative questions. Can we modify collective decision making procedures in ways that significantly improve outcomes, without adding too much in terms of complexity or implementation cost? The exercise in chapter 3 is a step in this direction. We amend usual voting procedures by allowing voters to choose the timing of their votes, and releasing interim summary information about vote tallies. This modification of rules does not seem too demanding given today's information technology. We show that if voters' fundamental objectives are closely aligned, such a modification allows information to be efficiently aggregated.

In the analysis of Chapters 2 and 3, do not allow explicit pre-play communication among the voters. This seems reasonable in some cases (e.g. in national elections), where the electorate is large and dispersed, but unrealistic in others (e.g. in jury decisions). Most elections will present *some* opportunity for non-binding pre-play communication, but it may be imperfect for reasons of time, cost, complexity or communication difficulties. The analytical framework adopted in this study may be a reasonable approximation for some of these scenarios. As mentioned, it would be nice to build a richer framework capable of handling more diverse preferences and opportunity for pre-play cheap talk, but this raises significant technical difficulties which the literature on the subject is far from resolving satisfactorily.

Essay 1: Optimal Voting Rules with Divergent Preferences

We investigate a model where the decision making is delegated to a group of individuals with disparate information but common preference. The group arrives at a decision using a voting mechanism where the voting rule is chosen by a designer who may have preference difference from that of group members. The designer faces a trade-off between choosing a rule that leads to maximum utilization of information and a rule that aggregates information in a way to suit her own preference. We show that the designer resolves this trade off by going to one of the extremes. In other words, either she chooses a rule that efficiently utilizes almost all available information or a rule that leads to almost all information being ignored. We provide a full characterization of these two cases based on the designer's preferences.

We show that if the designer and the group has opposing biases in favour of one of the alternatives, then the best voting rule for the designer is nearly the one that is best for the group. If, on the other hand, the designer has similar, but relatively stronger, bias as the group for one of the alternatives, the optimal rule is one which requires near-unanimity rule for the favoured alternative. Intuitively, this gives the group greater incentive to vote for the designer's favoured alternative uninformatively (since the group has similar bias, it will vote uninformatively in favour of the designer's preferred alternative). We conclude by noting that for a wide range of designer's preferences, the optimal voting rule is almost independent of the designer's preference. In other words, decisive power (votes) is more important than constitutional power in many of the scenarios considered.

Essay 2: Information Aggregation in Multi-Round Elections

In this essay we examine a model of sequential voting to examine its effects on the information aggregation properties of the voting rules. In contrast to some existing studies which claim that vote sequencing cannot change the equilibrium outcome of a voting game, we show that it does if the sequencing is endogenous. If the voters have common interest, or in other words if the sole purpose of the election is information aggregation, we show that with flexible timing there always exists a Bayesian Nash equilibrium, where information is fully aggregated. Moreover, if the model is symmetric, full information aggregation can be achieved as a symmetric equilibrium outcome. One of the most important features of our result is that this efficient equilibrium can be obtained irrespective of the voting cutoff.

In the common interest case, if pre-play communication among the voters is allowed, all information is revealed instantly, and full information aggregation is achieved. The sequential voting procedure with flexibility in timing of the votes achieves the same in the absence of communication. A natural question is how well this indirect mechanism works in the presence of conflict of interest among the voters. No voting procedure will generally achieve full-information aggregation in this case since the voters have an incentive to strategically withhold their private information. We illustrate, with a two person example with conflicting preferences, that the equilibria of the sequential voting game are outcome equivalent to the set of efficient equilibria of the same voting game with pre-play communication. Unfortunately, we do not have a result that establishes a direct connection between the above two sets of equilibria under a more general environment.

Essay 3: Endogenous Information and Committee Design

The third essay of the thesis explores the moral hazard problem of individuals in the context of committee design with endogenous information. We identify two kinds of effects on the incentive for collecting information that could arise as a result of a larger committee size. Since information is a public good, with endogenous information collection there could arise a *free-rider effect* leading to a fall in the quality of information collected by each individual when several members are entrusted with the task. However, there could arise a *complementarity effect* as, in some situations, one's information may not be valuable in itself but becomes valuable in conjunction with others' information. The second effect bolsters the incentive to collect information in a larger committee and could lead to better quality information collected by each member.

In a model involving different levels of quality of information, the designer could face a trade-off between quantity and quality of information in some situations, while deciding on the size of the committee. We show that the resolution to this trade-off can go either way. We identify parameter zones where a smaller committee performs strictly better and thus provide a strict violation of one of the Condorcet jury theorems.

2. Optimal Voting Rules with Divergent Preferences

2.1 Introduction

Often we encounter situations where the decision maker or the society has to delegate the decision making to a group of agents (such as a jury or a committee or a legislature). With exogenous information, a group has access to greater information than a single individual and hence in an environment with imperfect information has an obvious advantage in decision-making. Differences of opinions among the member of this group may arise due to different information, which then can be aggregated to arrive at a group decision using some social choice rule.

One problem the group decision making may encounter is that of aggregating the disparate information of individual members. Not all social choice rules perform equally well in this respect. A social choice rule is any rule that maps individual opinions into a group decision. If there is no inherent conflict of interest among the members of the group, the best social choice rule for the group is the one that aggregates individual information most efficiently. This is well documented in the literature on social choice and voting[1]. If the designer of the choice rule has the same preference as the group, the efficient rule for the group is the most efficient one for the designer. But if the designer's preference differs from that of the group, the optimal rule for the group may not be optimal for the designer because of the obvious reason that the designer has an incentive to distort the choice rule to suit her own preference.

In many decision problems, time inconsistency of preference is well recognized. At the outset, an optimal plan for future course of actions can be identified, but when the time comes to execute these actions, they are no longer preferred. Time inconsistency can be observed in every sphere of life - from individual decisions to government policy making or societal response to an event. Individuals who have decided to quit smoking do it. Couples who have pledged to a future plan of savings over a life span spend money on a top end car by deviating from the original plan. Government policies are even more prone to time inconsistencies. Populism in the face of election, giving in to pressures from interest groups are commonly observed phenomena. In the sphere of law, long term justice demands impassionate reasoning. But in the emotionally charged atmosphere of a criminal trial, the jury's decision often gets affected by sympathy and emotions rather than pure logic. For example, in a trial for a heinous crime committed to a child, a jury may be more trigger happy, while the long term view of natural justice demands stronger evidence for conviction. Time inconsistency may arise due to strategic reasons as well. In a situation involving strategic interactions among multiple agents, an agent may make long term gains by committing

[1]Starting from Condorcet [12] in the eighteenth century, there exists a large literature that deals with voting and elections as social choice mechanisms.

to a particular course of actions against others under certain circumstances, but ex post there may still exist incentives for deviation from the promised course of action for the sake of short term benefits.

To avoid the time inconsistency problems, two types of strategic responses are common. One is to engage in precommitments that reduce the opportunities to give in for temptations. Individuals commit to a generous pension plan to avoid the temptation of unnecessary current consumption. The other option available for avoiding time inconsistency problems is delegation of the decision making to some other agent or group of agents. In monetary and exchange rate policies, for instance, the credibility of the policy is very important and it can be costly in the long run to give in for temptations in the short run (Kydland and Prescott [32] and Barro and Gordon [4]). The time inconsistency problem in the context of monetary policy has been extensively studied in literature. Rogoff [48] showed that appointing an independent conservative central bank that put more weight on controlling inflation than increasing output would reduce the discretionary inflationary bias in monetary policy. The rules by which the central bank is governed are set by the government while the actual decision is taken by the bank independently. In the context of US trade policy, the incumbent President has very little discretion. There exist mandated sanctions by law implemented by US Trade Representatives against certain kind of behaviour of the trading partners of the United States. The delegation of decision is made to make the commitment to sanctions credible, thus avoiding the potential time inconsistency problems. Both precommitment and delegation of decision-making to other agents can be viewed as institutional devices to improve or eliminate the time inconsistency problems.

The purpose of this discussion is to argue that there may arise many real world situations where long term interest of the society may be in conflict with its short term decision makers. Often it is the case that the society sets the rules of decision making while the power of decision making rests with others. In this chapter, we analyze a model of a social institution where decision making is delegated to a group of agents. The said group of agents, who may or may not have the same preference as the society, arrives at a decision using some collective choice mechanism set by the society. Many social institutions fit into this category. Legislative decisions and jury trials are two of the more prominent ones.

Voting is one of the most commonly observed collective choice mechanisms. Hence, it is interesting to examine a situation like above in the context of a simple voting model where the designer chooses the voting rule and the group takes the decision. In the presence of conflict of interest between the designer and the group, the designer faces a trade-off while choosing the voting rule. She may distort the voting rule from that which leads to efficient information aggregation to suit her own preference, but in the process loses valuable information. We show that the resolution of this trade-off has the nature of a corner solution. The best solution for the designer is either to induce usage (almost) all the information, or to induce the group to take the decision using as little information as possible in the decision making. The intuition behind this result lies in the convexity in the value of information[2]

[2] The convexity in the value of information has been first identified by Radner and Stiglitz[47] though in a different model. But the result is general and holds under most environments.

arising out of the complementarity of the private information possessed by different members of the group. It is well known that value of information is convex in a decision problem. As it turns out in our model, even when the decision is delegated, the convexity property carries over.

We use a two alternative binary signal model where the group chooses one of the alternatives using a voting rule. Each member of the group, after observing her private signal, casts a vote for one of the alternatives. The votes are aggregated using the voting rule chosen by the designer. This is a standard model used extensively in the literature on sophisticated (or strategic) voting[3]. We show that for a wide range of values of the designer's preference parameter, the optimal voting rule for the designer is almost the same as the optimal voting rule for the group. If the group members vote in a sophisticated manner, then they try to rectify any deviation from their optimal voting rule by altering their equilibrium behaviour. Hence, the change in the voting rule by the designer to suit her own preference does not have the desired effect (or at least to the extent she desired) while resulting in loss of valuable information.

More specifically, our results show that if the designer and the group members have opposing biases[4], the optimal voting rule for the designer is almost[5] the same as that of the group. On the other hand, if the designer has a relatively stronger but similar bias as the group, then a unanimity or

[3]See Austen-Smith and Banks[3], Feddersen and Pesendorfer[21, 22], Miller[43] to name a few among many.

[4]An agent is biased in favour of of an alternative, if in the absence of any information she chooses that alternative.

[5]By almost same here we mean that the designer's optimal voting rule is either same as the group's optimal rule or just one off in either direction.

near-unanimity rule for the favoured alternative is optimal for the designer. Both results betray the initial intuition. With opposing biases, one would think that the designer will have greater incentive to manipulate the voting rule. While in case the designer has very strong bias in favour of one of the alternatives, it is likely to be the case that she will have incentive to make it easier for the group to choose that alternative. But exactly the opposite happens because of the sophisticated behaviour of the members, which the designer takes into account. The last result is somewhat related to what Feddersen and Pesendorfer[22] have shown. In their case, the probability of choosing an alternative wrongly may be higher under unanimity rule for that alternative than under some interior rule. We obtain a similar type of result though the equilibrium they select is different from ours.

While a significant number of studies in the literature on strategic voting have dealt with properties of different voting rules[6], most are concerned with finding the rule that aggregates information in the best possible way[7]. Austen-Smith and Banks [3], in a very important contribution to this field, have shown that with a binary signal structure there exist a unique voting rule that aggregates information efficiently and hence maximizes the group's ex-ante pay-off. We use this result extensively throughout in our model. Feddersen and Pesendorfer [21, 22], Duggan and Martinelli [20], Martinelli

[6]Some papers have examined the efficiency of different voting rules under the assumption of *sincere voting.* Sincere voting holds when voters vote the way they would had they been called upon to make the decision alone. See Miller [43], Ladha [33] and Berend and Paroush [6] for analysis of this. Young [53, 54] analyzes the same with more than two alternatives. We do not assume sincere voting a-priori, though this may arise as a result of equilibrium behaviour.

[7]With a signal structure that allows for more than two realizations, generally there does not exist any voting rule that aggregates information efficiently. See Austen-Smith and Banks [3].

[39], Meirowitz [42] and Wit [51] discussed asymptotic properties of different voting rules under various environments[8].

In an important contribution, McLennan [41] has shown that if the group members have common interest, the strategy profile that maximizes the ex-ante (common) pay-off function given the voting rule constitutes a Bayesian Nash equilibrium. Since this equilibrium is the ex-ante pay-off maximizer for the group, this equilibrium is optimal among the class of Bayesian Nash equilibria from the point of view of the group. Chakraborty and Ghosh [10] have shown that there always exists an efficient equilibrium in pure strategies and they have provided a complete characterization of the pure strategy efficient equilibrium for a voting game in the binary signal case. In our model, we choose this equilibrium for the purpose of comparison across different voting rules.

One important aspect of our model that merits some discussion here is that we don't allow for direct communication among the group members. With common interest among the group members, communication prior to voting would lead to all information becoming public knowledge. Hence, the decision will be independent of the voting rule, with every individual voting one way or the other. The designer can then do nothing by manipulating the voting rule. We show that even in the absence of communication the designer can do very little given sophisticated voting by the voters.

Our model can also be related to the literature on committee design and delegation. One strand of literature in this area extends the information transmission model of Crawford and Sobel [14] to examine how a decision-

[8]Some have dealt with richer signal spaces, while others have incorporated heterogeneity among jury's preferences.

maker's welfare can be influenced by varying the composition of a committee consisting of members with useful information ("expertise") whose policy preferences differ from that of the decision-maker (or the organization). Gilligan and Krehbiel [26], Austen-Smith [1, 2], Krishna and Morgan [30, 31], Battaglini [5] examine situations where all committee members have the same information but may have different preferences. Wolinsky [52] examines a model where experts have the same preference but receive different signals, but the power of decision making rests with the designer, while in our model the decision is delegated[9] to the group members. Li and Suen [37], Dessein [17] and de Garidel-Thoron and Ottaviani [15] have studied models of delegation; but while the focus of these studies are mostly on composition of the group, we focus on the choice of the decision rule for a group with given preference.

Other notable contributions to the literature on voting with private information are the following. Dekel and Piccione [16] analyze sequential, rather than simultaneous, voting procedures. Persico [46] studies a voting model with endogenous information to compare among different voting rules in terms of the incentives generated for information acquisition, as well as their information aggregation properties. Coughlan [13], Doraszelski, Gerardi and Squintini [19] and Gerardi and Yariv[23] study voting behaviour when voters can communicate. While these are interesting issues in themselves, our focus lies elsewhere in the current exercise.

The rest of the chapter is organized as follows. Section 2 analyzes the model and states the results. Section 3 concludes.

[9]For a general model on the theory of delegation see Holmstrom [28].

2.2 The Model and Results

We consider a model of collective choice with $n+1$ agents where there are two alternatives to choose from. We will use the terminology of the jury models for the rest of the chapter. Following the common practice in the jury literature, we assume that the collective decision must be made about whether to convict or acquit a defendant. The two social alternatives are — C (Conviction) or A (Acquittal). There are two possible states of nature - G (Guilty) and I (Innocent). All the agents would like to convict the guilty and acquit the innocent under complete information. With incomplete information, two types of error may arise as a result of the decision - namely convicting the innocent and acquitting the guilty. The utility function for agent i, $u^i(d,\omega)$, where $d \in \{C, A\}$ represents the ultimate decision and $\omega \in \{G, I\}$ is the state-of-the-world, is the following:

$$u^i(d,\omega) = \begin{cases} -q_i & \text{if } d = C, \omega = I \\ -(1-q_i) & \text{if } d = A, \omega = G \\ 0 & \text{otherwise.} \end{cases} \quad . \tag{2.1}$$

for $i \in \{0, 1, 2..., n\}$ where $q_i \in (0,1)$.

Agents $1, 2, ..., n$ are referred to as jurors and they are collectively referred to as the jury, J. The members of the jury share a common preference, that is $q_i = q_J$ for all $i = 1, 2, ..., n$. Each member of the jury gets a signal on the actual state of nature. We consider a binary signal structure. Each signal has two possible realizations: "guilty" (g) or "innocent" (i). The signals are partially informative. For each $j \in J$, $\Pr(g|G) = \Pr(i|I) = p$ with $p > \frac{1}{2}$.

With probability greater than $\frac{1}{2}$, each signal is *correct.* We also assume the signals to be conditionally independent. The common prior probability on the event $\omega = G$ is π for all agents.

Generally speaking, the problem can be posed as a standard mechanism design problem. Type of juror j, t_j, is private information and belongs to the type space $T_j = \{i, g\}$. Each juror sends a message to the mechanism designer and each juror's message is a function of his type, t_j, i.e. $m_j : T_j \longrightarrow M_j$ where M_j is the message space for juror j. The mechanism designer sets a choice rule, F that maps the vector of messages into a decision. In other words, let $M = M_1 \times M_2 \times \cdots \times M_n$. Then, $F : M \longrightarrow \{C, A\}$. Given any F, each juror chooses her message to maximize her pay-off. The designer chooses F to suit her own preference.

In the ongoing exercise, we refrain from analyzing the general mechanism design problem. We restrict ourselves to a particular class of mechanism - namely the voting mechanism with single transferable vote. Each juror can cast only one vote in favour of one of the alternatives and then the decision is taken according to some monotonic or cutoff rule. In the language of mechanism design, we we restrict the message space for each juror to be binary, i.e. $M_j = \{C, A\}$ and the choice rule to be monotonic in the number of messages of each type, i.e. if the decision is conviction for some number of C votes, it must be the same for any higher number. We also restrict ourselves to deterministic mechanisms only.

Why do we analyze such a restrictive mechanism? Voting is one of the age-old practical mechanisms using which collective choice problems are resolved. One advantage of a voting mechanism is that it is easily implementable. Moreover, a voting mechanism with a single transferable vote has an ethical flavour and is often seen as a desired democratic practice. Thus it is commonly observed in real world and demands closer attention in its own merit. Social scientists are aware of its importance. Starting from Condorcet' in the eighteenth century, economists and political scientists have contributed significantly in analyzing voting mechanisms. Even in recent times, a large number of studies, using sophisticated mathematical techniques have discussed relative merits and demerits of different voting rules[10] in optimizing the social outcome (Austen-Smith and Banks [3]Feddersen and Pesendorfer [21, 22], Duggan and Martinelli [20], Martinelli [39], Meirowitz [42] and Wit [51]). Our attempt here is to contribute towards this growing literature when there is conflict of interest between the voters and the designer of the voting rule.

We assume that the joint decision is taken by the jury J by playing a non-cooperative simultaneous move Bayesian game. After observing their private signals, each juror $j \in J$ takes an action $x_j \in \{C, A\}$ simultaneously and independently of other jurors. The action $x_j = C$ is interpreted as a vote for conviction while $x_j = A$ is a vote for acquittal. The individual votes are then aggregated into a decision d by a voting rule which is chosen by agent 0. Henceforth, we will refer to agent 0 as the "designer" with a preference

[10] One can think of more complicated voting rules rather than the simple cutoff rule we consider here. Borda count is one such rule where alternatives are ranked by the voters in order of preference and then the winner is chosen by summing over the ranks obtained by each alternative. In our case, however, this does not have any significance since we are in an environment with two alternatives. Scoring rule is another rule where each alternative is given a score instead of a vote by each voter and then aggregate scores are used to pick the winner. We however restrict ourselves to cutoff voting rules with single non-transferable votes

parameter q_D which may or may not be different from q_J. Both q_J and q_D are common knowledge.

The timing of the game is as follows. The designer first chooses the voting rule. The jurors then play the voting game. The votes are then aggregated and the decision is reached according to the specified voting rule.

Let the voting rule R denote the minimum number of C votes required to convict. Given any R, when all voters vote informatively the jury's ex-ante payoff is the following:

$$V(R, n; q_J) = -\sum_{x=0}^{R-1} \binom{n}{x} p^x (1-p)^{n-x} \pi (1-q_J) - \sum_{x=R}^{n} \binom{n}{x} p^{n-x} (1-p)^x (1-\pi) q_J. \tag{2.2}$$

Suppose we denote $R_s(n) = \arg\max_R V(R, n; q_J)$. We also assume that $\frac{\pi(1-q_J)}{(1-\pi)q_J}$ is not too extreme in either direction such that $R_s(n) \in \{1, 2, ..., n\}$. As Austen-Smith and Banks[3] have pointed out, $R_s(n)$ is the unique rule under which all voters vote informatively in equilibrium. Informative voting by a juror means that a juror votes for C if and only if he receives a signal g. The intuition behind this result is that since $R_s(n)$ is optimal, the jurors want to move the voting rule towards $R_s(n)$. Suppose $R < R_s(n)$. Then when all others vote informatively, if a juror votes informatively the probability of conviction is higher than what is optimal. Hence the juror wants to rectify this by voting A irrespective of signal. A similar argument holds for $R > R_s(n)$ as well.

Henceforth, we will call $R_s(n)$ the *jury-optimal rule.* Obviously, since under the jury-optimal rule every juror votes informatively, information is fully aggregated and the ex-ante pay-off of the jury is maximized. Now if the designer's preference is the same as the jury's, the jury-optimal rule is optimal for the designer as well. But if the designer's preference is different from the jury's preference, the designer may have an incentive to distort the voting rule to suit her own preference. Apparently, the pay-off maximizing voting rule for the designer should be different from the jury-optimal rule which maximizes the jury's pay-off. We show that, quite surprisingly, the designer's optimal voting rule is not very different from the jurors for a very wide range of preferences of the designer.

Before we proceed further, we consider the benchmark case where all private signals become common knowledge. For the jury J of size n, consider the decision problem when the signal vector $s = (s_1, s_2, ..., s_n)$ is common knowledge. In this case, from the jury's point of view, the optimal decision rule is to choose C whenever $\Pr(G|s) > q_J$ and to choose A if $\Pr(G|s) < q_J$. With private signals, whether a jury would be able to implement the full information decision rule above depends, along with other things, on the voting rule R. The following definition characterizes the voting rules that satisfy full-informational equivalence.

Definition 2.1. *A voting rule R satisfies full-information equivalence for a jury of size n if there exists a Bayesian Nash Equilibrium σ_J such that* $\Pr(d_R = C|\sigma_J, s) = 1$ *whenever* $\Pr(G|s) > q_J$ *and* $\Pr(d_R = C|\sigma_J, s) = 0$ *whenever* $\Pr(G|s) < q_J$, *where d_R is the decision under voting rule R and σ_J is the equilibrium strategy profile for jury J.*

We now establish a property of the jury-optimal rule that we will extensively use in what follows. The jury-optimal rule under which every juror votes informatively in equilibrium satisfies full-information equivalence.

Lemma 2.1. *The jury-optimal rule, $R_s(n)$, for a jury J of size n is the unique voting rule that satisfies full information equivalence.*

Proof. See Austen-Smith and Banks [3]. ∎

Since the jury-optimal rule satisfies full-information equivalence, the outcome of a voting game with n jurors and voting rule $R_s(n)$ is the same as the decision problem where an individual juror has access to all n signals. The optimal decision for the individual juror will then be to choose C if the number of guilty signals out of n is greater than or equal to $R_s(n)$. If the jury's relative cost of convicting an innocent defendant, q_J, falls, the jury would require less evidence to convict. Hence, if private signals become common knowledge, the number of guilty signals out of n required to convict falls with q_J. Therefore, for a fixed jury of size n, $R_s(n)$ falls with q_J.

When the size of the jury increases, more signals are available. Hence, the optimal number of guilty signals required to convict changes for a given preference of the jury. In other words, $R_s(n)$ changes with n. In the next lemma we show that $R_s(n)$ changes at half the rate of change in n.

Lemma 2.2. *For any n, $R_s(n+2) = R_s(n) + 1$.*

Proof. Take any n. Since $R_s(n)$ satisfies full-information equivalence

$$\left(\frac{p}{1-p}\right)^{n-2R_s(n)+2} > \frac{\pi(1-q_J)}{(1-\pi)q_J} \geq \left(\frac{p}{1-p}\right)^{n-2R_s(n)}.$$

Similarly,

$$\left(\frac{p}{1-p}\right)^{(n+2)-2R_s(n+2)+2} > \frac{\pi(1-q_J)}{(1-\pi)q_J} \geq \left(\frac{p}{1-p}\right)^{(n+2)-2R_s(n+2)}.$$

By Lemma 2.1, we know that $R_s(n)$ satisfies full information equivalence for all n. Now for all n, both set of inequalities are satisfied if and only if $R_s(n+2) = R_s(n) + 1$. ∎

So far we have restricted ourselves to characterization of the jury-optimal rule. We now turn our attention to the efficient equilibrium under any voting rule. Let $\sigma_i(s_i)$ denote the probability of voting C for juror i after receiving the signal s_i. We say that a juror votes informatively in any pure strategy equilibrium if $\sigma_i(g) = 1$ and $\sigma_i(i) = 0$. Given a jury of size n and the voting rule R, following McLennan [41], we know that there exists an efficient pure strategy equilibrium. Chakraborty and Ghosh [10] have characterized this efficient equilibrium. Generally, for any arbitrary voting rule, not all voters vote informatively in this equilibrium. Only a subset of voters vote informatively. The following proposition characterizes the efficient pure strategy equilibrium.

Proposition 2.1. *(Chakraborty and Ghosh [10], Persico [46]) Assume there are n informed jurors. Then, if $R < R_s(n)$, at the most efficient equilibrium in pure strategies, a number $n - m_R \leq n$ of jurors vote* Acquit *regardless of their signal, and the remaining m_R jurors vote informatively where m_R is such that $R_s(m_R) = R$. If $R > R_s(n)$, at the most efficient equilibrium in pure strategies, a number $n - m_R \leq n$ of jurors vote* Convict *regardless of their signals, and the remaining m_R jurors vote informatively where m_R is such that $R_s(m_R) + n - m_R = R$.*

Proof. See Chakraborty and Ghosh [10]. ∎

Notice that given any voting rule there exist numerous Bayesian Nash equilibria in the voting game with n jurors. The objective of this exercise is to characterize the optimal voting rule from the designer's perspective when the preference of the designer differs from that of the jury. There is an equilibrium selection problem here. An obvious choice is to select the most efficient equilibrium given any voting rule. We assume that the jurors are sophisticated enough to play the efficient equilibrium[11].

In the efficient equilibrium, not all jurors vote informatively. But those who vote informatively make sure that all the information used in equilibrium is efficiently aggregated. If m_R jurors vote informatively in the efficient equilibrium given voting rule R, then the outcome in this efficient equilibrium is equivalent to the outcome of a voting game with jury of size m_R with the voting rule $R_s(m_R)$. Since jury-optimal rule efficiently aggregates information, this is outcome equivalent to the decision problem where the signals of the m_R jurors are common knowledge.

The above argument shows that in the efficient equilibrium, the jurors adjust their equilibrium behaviour to offset the effect of a change in voting rule to suit their preference. Hence, from the designer's point of view, manipulating the voting rule amounts to choosing a jury of size $m < n$, with the effective voting rule $R_s(m)$. By Lemma 2.1, this is equivalent to choosing m number of signals when the decision is taken optimally according to the jury's preference. The designer's problem then boils down to finding a feasible $m \leq n$ that maximizes her own ex-ante pay-off.

Since we assume that the jury plays the most efficient equilibrium given any voting rule, by manipulating the voting rule from the jury-optimal rule,

[11]There is a coordination problem here. But we ignore that for the sake of comparison across equilibria of similar nature.

the designer effectively makes less information count towards the ultimate decision. As long as there is a possibility that the potential aggregate pool of information is decision relevant for the designer, this is a net loss. On the other hand, distorting the voting rule to suit her preference can only have a limited success, since the jurors vote strategically to suit their preference in the efficient equilibrium. On balance, the decrease in value resulting from the loss of information outweighs the gain from distortion in the voting rule for a wide range of preferences for the designer.

A little discussion regarding the intuition behind Proposition 2.1 may be worthwhile here. If $R > R_s(n)$, then, in equilibrium, m_R is such that it satisfies $R_s(m_R) + n - m_R = R$. If the voting rule is biased in favour of A from the jury's point of view, some jurors vote uninformatively for C to rectify it. But this has to be done in a way such that the rest have exact incentives to vote informatively. That is possible if and only if the effective voting rule among the informative voters is the jury-optimal rule. On the other hand, if $R < R_s(n)$, m_R must satisfy $R_s(m_R) = R$, since the rest vote uninformatively in favour of A. Since, in our model, the designer only has R as the instrument to manipulate the decision in her favour, the equilibrium behaviour of the jurors puts bounds on the effective number of informative voters she can choose.

We now turn our attention to the designer's payoff. Notice that our discussion so far has shown that by changing the voting rule effectively the designer can choose a smaller jury which votes informatively after strategically adjusting the voting rule to suit its preference. In some cases, the designer can induce completely uninformative voting by the jury in favour of one of the alternatives. In other cases, she cannot induce uninformative voting by all. We denote by $V(m; q_D)$ the designer's ex-ante payoff when a $m-$member jury decides the outcome on the basis of voting rule $R_s(m)$. We now characterize an important property of $V(m; q_D)$ which we need to prove our main result.

Lemma 2.3. *Suppose $n > 2$. For any $m \leq n$, $V(m; q_D) > V(m-2; q_D) \Rightarrow V(m+2; q_D) > V(m; q_D)$.*

Proof. See Appendix A.1. ∎

The last lemma also implies that $V(m+2; q_D) \leq V(m; q_D) \Rightarrow V(m; q_D) \leq V(m-2; q_D)$. Now suppose that the designer has to choose the size of the jury, m, from the set $N = \{n_0, n_0+1, ..., n\}$ with $n_0 > 0$ where the decision is taken by the jury using the jury-optimal rule. As shown in the above lemma, the designer would effectively choose one of four numbers $\{n_0, n_0+1, n-1, n\}$ to maximize her ex-ante pay-off.

There are two types of errors that may occur as a result of the jury's decision. One type of error arises because, for any finite jury, there is a positive probability of convicting an innocent defendant. The other type of error arises because of the possibility of acquitting a guilty defendant. The preference parameter q_t, $t = J$ or D captures the relative weights the agents attach to these two types of errors. Let $\rho_t = \frac{\pi(1-q_t)}{(1-\pi)q_t}$ be the relative cost of acquitting a guilty defendant for an agent of type t. This parameter ρ_t indicates the effective preference of the agents. Intuitively, if $\rho_t > 1$, the agent is relatively more biased in favour of conviction which means he attaches more importance on the error arising from acquitting a guilty defendant. We will state our comparative static results in terms of ρ_D and ρ_J.

The preference parameter for the jury, ρ_J, determines the smallest size of the jury that will vote informatively. If ρ_J falls outside the interval $\left(\left(\frac{1-p}{p}\right)^n, \left(\frac{p}{1-p}\right)^n\right)$ for some n, then any jury of size n or less will vote uninformatively for either C or A, depending on which side of the interval ρ_J lies in. This happens because here all n signals taken together do not have any decision relevance for the jury. Hence, given ρ_J, there is a lower bound on the size of an informative jury[12].

We now show that the designer will either prefer a jury which takes the decision completely uninformatively[13] or one which almost fully utilizes the available pool of information. Let m denote the size of the jury that votes informatively. $m = 0$ indicates a jury that takes the decision uninformatively.

Lemma 2.4. *The designer's pay-off is maximized either at $m = 0$ or at $m = n - 1$ or n.*

Proof. See Appendix A.1. ∎

Lemma 2.4 provides some important insights towards understanding the nature of the solution for designer's problem. Either the designer wants to tap (almost) the total pool of available information or she does not want to utilize any information at all. This is a direct consequence of Lemma 2.3 which essentially proves that the designer's objective function is convex in

[12]This lower bound can be equal to 1. If $\rho_J = 1$, then even a single individual will have an incentive to vote informatively.

[13]Notice that as we have mentioned the designer only has one instrument - the voting rule - to affect the effective size of the jury. Whether she would be able to induce all the jurors to vote uninformatively using this instrument is a different issue that we will address later.

the number of signals. But, in our set up, the designer cannot choose the jury (or the number of signals utilized in the process of decision-making). She has only one instrument - the voting rule - which she can use only to affect the equilibrium behaviour of the jury.

For ease of exposition of our main result, we will call any voting rule that is the same as the jury-optimal rule, or one off in either direction, a *near jury-optimal rule.* In similar vein, we will call a voting rule *near unanimity rule* if it is the unanimity rule or one off from the unanimity rule. We show that the designer's optimal choice of voting rule is either a near jury-optimal rule or a near unanimity rule. We now formally define these two rules.

Definition 2.2. *Given any n, a voting rule R is a near jury-optimal rule if* $R \in \{R_s(n) - 1, R_s(n), R_s(n) + 1\}$.

Definition 2.3. *Given any n, a voting rule R is a near unanimity rule for* Conviction *if* $R \in \{n - 1, n\}$. *Similarly, a voting rule R is a near unanimity rule for* Acquittal *if* $R \in \{1, 2\}$.

By varying the voting rule, the designer can affect the number of jurors who vote informatively in the efficient equilibrium. Our next lemma states to what degree the designer can affect this number and how it depends on the jury's preference.

Lemma 2.5. *If* $\rho_J \in \left[\frac{1-p}{p}, \frac{p}{1-p}\right]$, *there does not exist any voting rule* $R \in \{1, 2, ..., n\}$ *that can induce completely uninformative voting by all jurors in the efficient equilibrium. For any other value of* ρ_J, *there exist at least one voting rule* $R \in \{1, 2, ..., n\}$ *where all the jurors vote uninformatively in the efficient equilibrium.*

Proof. For proving the first part of the lemma, we have to check that, at $R = 1$ and at $R = n$, at least one individual has incentive to vote informatively. If $\rho_J \in \left[\frac{1-p}{p}, \frac{p}{1-p}\right]$, then $R_s(1) = 1$. Hence, from Proposition 2.1, at $R = 1$, a single juror has incentive to vote informatively even when all others vote uninformatively for A in the efficient equilibrium. Now consider $R = n$. Suppose $n-1$ jurors vote uninformatively for C. That the other juror has incentive to vote informatively can be seen from $R_s(1) + (n-1) = n$.

For the second part of the lemma, it can be easily shown that for $\rho_J > \frac{p}{1-p}$, $R = n$ will induce completely uninformative voting by all jurors and for $\rho_J < \frac{1-p}{p}$, $R = 1$ will do the same. ∎

We will now state our main result in the next proposition. The proposition characterizes the efficient voting rule for the designer for different values of her preference parameter.

Proposition 2.2. *Consider a n-member jury with* $\rho_J \in \left(\left(\frac{1-p}{p}\right)^n, \left(\frac{p}{1-p}\right)^n\right)$ *which decides the outcome by playing the efficient pure strategy equilibrium under any voting rule.*

1. *For all* $\rho_J > 1$*, there exists a* $\bar{\rho}(\rho_J) > \rho_J$ *such that for* $\rho_D < \bar{\rho}(\rho_J)$ *the designer's optimal voting rule is a near jury-optimal rule and for* $\rho_D \geq \bar{\rho}(\rho_J)$ *the designer's optimal voting rule is a near unanimity rule for Conviction.*

2. *For all* $\rho_J < 1$*, there exists a* $\underline{\rho}(\rho_J) < \rho_J$ *such that for* $\rho_D > \underline{\rho}(\rho_J)$ *the designer's optimal voting rule is a near jury-optimal rule and for* $\rho_D \leq \underline{\rho}(\rho_J)$ *the designer's optimal voting rule is a near unanimity rule for Acquittal.*

Proof. We will prove the first part of the proposition. The proof for the second part is similar and hence omitted.

Since $\rho_J > 1$, an uninformative decision by jury is C and hence $V(0; q_D) = -(1-\pi)\, q_D$. Notice that if $\rho_D \leq 1$ the designer will always prefer an informative decision to an uninformative one. By Lemma 2.4, then the designer would prefer an informative jury of size n or $n-1$. Since the designer's relative cost of convicting the innocent defendant is greater than that of the jury, the designer will prefer a relatively higher threshold of conviction than the jury. If $R_s(n-1) = R_s(n)$, the threshold for conviction in an informative jury is stronger for the size $n-1$ than for size n, since the jury of sizes $n-1$ and n need the same number of g signals for conviction. Hence the designer will prefer a jury of size $n-1$. By similar reason, if $R_s(n-1) = R_s(n) - 1$, the designer will prefer a jury of size n.

Now consider $\rho_D > 1$. We first consider the case where $R_s(n) = R_s(n-1)$. Notice that for the jury

$$\left(\frac{p}{1-p}\right)^{n-2R_s(n)} \leq \rho_J < \left(\frac{p}{1-p}\right)^{n-2R_s(n)+2}.$$

It can be easily verified that $V(n; q_D) \geq V(n-1; q_D)$ if and only if $\rho_D \geq \left(\frac{p}{1-p}\right)^{n-2R_s(n)}$ in this case. For higher values of ρ_D, the designer is relatively more concerned about convicting the innocent and hence would prefer an informative jury of size n to a jury of size $n-1$, since the threshold for conviction is relatively higher in a jury of size n than in a jury of size $n-1$ in this case. Now if the designer can induce completely uninformative voting by the jury in favour of conviction, then by Lemma 2.4, the optimal size of informative jury is either $n-1$ or 0 if $\rho_D < \left(\frac{p}{1-p}\right)^{n-2R_s(n)}$ and either n or 0 if $\rho_D \geq \left(\frac{p}{1-p}\right)^{n-2R_s(n)}$.

Suppose that the designer can induce completely uninformative voting by choice of voting rule. Lemma 2.5 identifies when this is indeed the case. Then, we can write

$$\begin{aligned}
& V(n;q_D) - V(0;q_D) \\
= & -\pi(1-q_D)\sum_{x=0}^{R_s(n)-1}\binom{n}{x}p^x(1-p)^{n-x} \\
& -(1-\pi)q_D\sum_{x=R_s(n)}^{n}\binom{n}{x}(1-p)^x p^{n-x} + (1-\pi)q_D \\
= & (1-\pi)q_D\sum_{x=0}^{R_s(n)-1}\binom{n}{x}(1-p)^x p^{n-x} \\
& -\pi(1-q_D)\sum_{x=0}^{R_s(n)-1}\binom{n}{x}p^x(1-p)^{n-x} > 0 \\
\Leftrightarrow & \rho_D = \frac{\pi(1-q_D)}{(1-\pi)q_D} < \frac{\sum_{x=0}^{R_s(n)-1}\binom{n}{x}(1-p)^x p^{n-x}}{\sum_{x=0}^{R_s(n)-1}\binom{n}{x}p^x(1-p)^{n-x}} = \rho_1(\rho_J)
\end{aligned}$$

That $\rho_1 > \left(\frac{p}{1-p}\right)^{n-2R_s(n)}$ can be easily verified by some simple algebra.

If the optimal size of the informative jury is n, the voting rule $R_s(n)$ can achieve it by Lemma 2.1. If it is $n-1$, then it can be achieved by choosing the voting rule $R_s(n)+1$. To see that this is indeed the case, verify that $R_s(n-1)+n-(n-1) = R_s(n)+1$ since $R_s(n-1) = R_s(n)$. If $\rho_J > \frac{p}{1-p}$, then by Lemma 2.5, $R = n$ will induce the jury to vote uninformatively in favour of C. Hence, $\bar{\rho}(\rho_J) = \rho_1$ will ensure that for $\rho < \bar{\rho}(\rho_J)$, optimal R from designer's perspective is either $R_s(n)$ or $R_s(n+1)$ and for $\rho \geq \bar{\rho}(\rho_J)$, optimal R is n.

Now consider the case when $\rho_J \leq \frac{p}{1-p}$. By Lemma 2.5, there does not exist any voting rule that can induce uninformative voting by all jurors in efficient equilibrium. If $\rho_D \geq 1$, $V(2;q_D) \geq V(1;q_D)$ is ensured from the fact $R_s(1) = R_s(2) = 1$ and the designer will prefer a relatively easier threshold for conviction. Hence, the designer's effective choice of the informative jury is limited to three numbers $\{2, n-1, n\}$. Now for any ρ_D, $V(2;q_D) \geq V(0;q_D) \Rightarrow \max\{V(n;q_D), V(n-1;q_D)\} \geq V(2;q_D)$ by Lemma 2.4. Hence, for any $\rho_D \leq \rho_1$, the optimal size for the informative jury is n or $n-1$. For $\rho_D > \rho_1$, the designer would be better off if she can induce uninformative voting by all members in favour of conviction, but the restriction on jury's preference makes sure that she cannot achieve this by manipulating the voting rule. Hence, the effective comparison for the designer is between choosing between an informative jury of size n and that of size 2. Now,

$$
\begin{aligned}
& V(n;q_D) - V(2;q_D) \\
= & \left[(1-\pi) q_D \sum_{x=0}^{R_s(n)-1} \binom{n}{x} (1-p)^x p^{n-x} - \pi (1-q_D) \sum_{x=0}^{R_s(n)-1} \binom{n}{x} p^x (1-p)^{n-x}\right] \\
& - \left[(1-\pi) q_D p^2 - \pi (1-q_D)(1-p)^2\right] \\
> & \ 0 \\
\Leftrightarrow & \ \rho_D < \frac{\sum_{x=0}^{R_s(n)-1} \binom{n}{x} (1-p)^x p^{n-x} - p^2}{\sum_{x=0}^{R_s(n)-1} \binom{n}{x} p^x (1-p)^{n-x} - (1-p)^2} = \rho_2(\rho_J).
\end{aligned}
$$

That $\rho_2 > \rho_1 > \left(\frac{p}{1-p}\right)^{n-2R_s(n)}$ can be verified easily for any ρ_J. We now choose

$$\bar{\rho}(\rho_J) = \begin{cases} \rho_2(\rho_J) & \text{if } 1 < \rho_J \le \frac{p}{1-p} \\ \rho_1(\rho_J) & \text{if } \rho_J > \frac{p}{1-p} \end{cases}.$$

We have already discussed that optimal voting rule for the case $\rho_J > \frac{p}{1-p}$. If $\rho_J \le \frac{p}{1-p}$, then optimal size of informative jury for the designer is n or $n-1$ for $\rho_D < \rho_2(\rho_J)$. Hence, the efficient voting rule from the designer's point of view is $R_s(n)$ or $R_s(n+1)$. For $\rho_D \ge \rho_2(\rho_J)$, optimally the designer wants two jurors to vote informatively under the constraint that she cannot induce complete uninformative voting. This can be implemented by choosing a voting rule $R = n-1$. To see this verify that $R_s(2) + (n-2) = n-1$ since $R_s(2) = 1$.

The proof for the case when $R_s(n-1) = R_s(n) - 1$ is similar except that ρ_1 and ρ_2 are now obtained by equating $V(n-1; q_D)$ with $V(0; q_D)$ and $V(2; q_D)$ respectively. The reason behind this is now a jury of size n has a relatively stronger threshold for conviction than a jury of size $n-1$ and hence for high values of ρ_D the designer would prefer the jury of size of $n-1$. Moreover, the designer can induce an informative jury of size $n-1$ with the voting rule $R_s(n) - 1$. From Proposition 2.1, one member then votes for acquittal uninformatively and the rest votes informatively. That this is the efficient equilibrium can be verified by $R_s(n-1) = R_s(n) - 1$.

Combining the above two cases, we complete the proof of the proposition. For $\rho_D \ge \bar{\rho}(\rho_J)$, the designer's optimal voting rule is a near unanimity rule and for $\rho_D < \bar{\rho}(\rho_J)$, the optimal voting rule is a near jury-optimal rule when $\rho_J > 1$. ∎

The meaning of the last proposition can be summarized as follows. If the designer's bias is either opposite that of the jury or similar but close, the designer is best off by not distorting the voting rule too much from what the jury itself would prefer. Specifically, she would find it optimal to choose exactly the jury's own optimal rule, or a rule which differs from it by the margin of a single vote. For example, suppose the jury has an acquittal bias (i.e. absent any information, the jury would prefer to acquit). If the designer has a conviction bias, or an acquittal bias of similar magnitude, the best thing for the designer to do is to set a rule that attempts to (nearly) maximize the amount of information being utilized in the jury's decision. This is achieved by choosing the voting rule that the jury most prefers. Our analysis shows that the jury will respond to any distortions in the rule by contracting the amount of information used in the decision in a way that (almost) neutralizes the distortion. The designer, therefore, fails to influence the way the utilized information is aggregated into the decision, but simply induces a loss of available information. She is better off avoiding this pure information loss.

The strategic considerations are dramatically different when the designer and the jurors have similar biases (say in favour of acquittal), but the designer's bias is much more extreme than that of the jury. For example, consider a situation where the designer ideally wants a much higher burden of proof to convict a defendant than the jury. In such situations, the designer may prefer to *minimize* rather than maximize the informativeness of the decision. Since, in the presence of little or no information, the jury leans the same way as the designer, she may want to exploit this feature and avoid a lot of convictions in marginal cases. It is interesting to observe what kind of rule achieves this. In the case of similar but very disparate biases (say both lean towards acquittal but the designer more strongly so), she can achieve informational shutdown by making acquittal very *difficult*, i.e. by choosing a rule that requires unanimity to acquit! Such a rule makes a single conviction vote too risky, and the jury responds by voting for acquittal, disregarding its information altogether.

It is interesting to contrast this result against that of Feddersen and Pesendorfer [22], who pointed out some undesirable properties of the unanimity rule in jury trials. In particular, they demonstrated that the indirect strategic effect of changing the conviction threshold may outweigh the direct effect, to produce counter-intuitive results. In some circumstances, it may be more likely that an innocent person is convicted when conviction requires unanimity rather than simple majority. Feddersen and Pesendorfer [22], however, do not conduct explicit welfare analysis, which is a deficiency since these problems involve a trade-off between two kinds of errors (convicting the innocent and acquitting the guilty) and an increase in one kind of error can be more than compensated in principle by a decrease in the other. This is especially important when the designer and the juror's preferences are in conflict, which may prompt the former to manipulate the rule to suit her own objectives.

Our results show that optimal manipulation is either (nearly) zero, i.e choosing the jury optimum rule, or maximal, i.e a near unanimity rule, which leads to the greatest information loss. Hence, unanimity *is* a salient rule in voting, in the sense that it is the preferred rule *whenever* a designer finds it optimal to manipulate the decision making process through her constitutional powers. A further surprising feature is that this happens in a direction that is diametrically opposite to what first intuition may suggest. Again, consider the example of a legislature which may place a lot of weight on the welfare loss from wrongful convictions, but may be worried about juries being relatively trigger happy, i.e. willing to convict on a lower burden of evidence. Feddersen et al suggested that raising the bar by requiring unanimity for conviction may be counter productive. Our results show that if manipulation is useful at all, the pendulum swings all the way to the other end, i.e. the optimal rule is to make conviction very easy, by requiring (near) unanimity for *acquittal*! Analogously, liberal minded constitutional designers (i.e. those receptive to change) may want to create highly conservative constitutions, requiring strong consensus to vote for change, and vice versa.

2.3 Concluding Remarks

We investigate a model where the decision making is delegated to a group of individuals with disparate information but common preference. The group arrives at a decision using a voting mechanism where the voting rule is chosen by a designer who may have conflict of interest with the group members. The designer faces a trade-off between choosing a rule that leads to efficient information aggregation and a rule that suits her own preference. We show that the designer resolves this trade off by going to one of the extremes. In other words, either she chooses a rule that almost fully aggregates information or a rule that uses almost no information. We provide a full characterization of these two cases based on designer's preferences.

We show that if the designer and the group have opposing biases in favour of one of the alternatives, then the best voting rule for the designer is nearly the one that is best for the group. If, on the other hand, the designer has similar, but relatively stronger, bias as the group for one of the alternatives, the optimal rule is one which requires near-unanimity for the favoured alternative. Intuitively, this gives the group greater incentive to vote for the designer's favoured alternative uninformatively (since the group has similar bias, it will vote in favour of the designer's preferred alternative when it is voting uninformatively). We conclude by noting that for a wide range of preferences for the designer, the optimal voting rule is almost independent of the designer's own preference.

3. Information Aggregation in Multi-round Elections

3.1 Introduction

Voting is the most common way in which social choice problems are resolved. However, if voters have noisy private information about the possible consequences of various choices, usual (simultaneous) voting procedures fail to aggregate that information efficiently. This is most clearly seen in a model in which voters have identical preferences but different information[14]. The decentralized voting outcome is usually not *informationally efficient*—welfare would increase if a central planner could (costlessly) collect everyone's information and implement a decision. Of course, this is only an ideal benchmark. In reality, for a planner to collect accurate information from thousands or millions of people is an exercise fraught with incentive and free-rider problems, to say nothing of administrative costs. A decentralized mechanism is, therefore, more or less a practical necessity. Hence, the search for better (yet practical) decentralized mechanisms is an important one for social choice

[14]It can be argued that generally the voters will have both private information as well as heterogeneity of preference. Nevertheless, the pure common interest case serves as a useful theoretical benchmark.

problems.

A significant number of theoretical and experimental studies, particularly in the context of jury trials, have tried to analyze this issue both in common interest elections and in the presence of some degree of heterogeneity in preferences. Ladha [33], Miller [43], Feddersen and Pesendorfer [21, 22], Duggan and Martinelli [20], Wit [51], Ladha, Miller and Oppenheimer [34] represent just a small sample of recent literature.

The problem of aggregating information arises because in a simultaneous voting two-alternative election, voters condition their votes on being pivotal given the voting rule. Austen-Smith and Banks[3] show that *sincere voting*[15] is not an equilibrium behaviour in general. Although for a binary signal structure, full informational efficiency can be achieved in a Nash equilibrium by properly choosing the voting rule, this is not true for a richer signal space. In a binary signal case, sincere votes can reflect the realized signals[16] of the electorate and, if the voting rule is optimally chosen, one can achieve the central planners outcome. But this is not true when the signal space is rich enough so that voters cannot mimic their signals just by using their votes.

The obvious way to resolve this problem is to introduce pre-play communication by allowing players to exchange messages prior to actual voting. With perfectly aligned (or even almost common) interests, all private information will be revealed, and then players will vote unanimously to achieve the informationally best outcome. But communication may not be feasible in all elections. Moreover, to achieve informational efficiency through com-

[15] Sincere voting is the voting behaviour in which each voter chooses the alternative yielding the highest expected payoff conditional on her private signal.

[16] We consider the case when sincere voting is informative, i.e. each players vote depends on the private signal.

munication, the message space has to be sufficiently rich so that players can convey their signals by their messages. Arbitrary communication mechanisms cannot achieve this end. Coughlan [13] examined a straw vote as a means of communication and showed that full informational efficiency can be achieved for any voting rule. But again this result holds only for binary signals and it fails as soon as the signal structure is more complicated. Doraszelski, Gerardi and Squintani [19] and Gerardi and Yariv [23] study voting behaviour when voters can communicate using a sufficiently rich message space which allows the voters to imitate their signals. But direct pre-play communication is often costly, particularly if the size of the electorate is large enough. Moreover, any mechanism that attempts to resolve a collective choice problem has greater appeal if it is indirect and easily implementable. Full pre-play communication between the players even in a medium size election is often too much to ask for.

What happens if direct communication is not feasible before voting? Chakraborty and Ghosh [10] showed that for conditionally independent signals full information equivalence can be achieved for any voting rule that allows divisibility of individual votes. This chapter attempts to provide an alternative mechanism that can improve the voting outcome significantly.

At first glance, it seems that instead of simultaneous votes, some kind of sequencing of the voting procedure could lead to better results, by allowing later voters to draw inferences from the votes exercised by those who preceded them, and thereby vote in a more informed way themselves. However, Dekel and Piccione [16] have shown that this intuition is generally not true. Under sequential voting, the set of equilibrium outcomes remains the same as under simultaneous voting. Callander [9] examines a model where voters obtain positive utility from voting for the winning candidate and finds that superiority of the voting procedure depends on the value of the common prior. Morton and Williams [44] compare between simultaneous and sequential voting outcomes empirically. All of these papers, however, consider rules that require the sequence in which the electorate must vote to be exogenous. This undoubtedly has practical relevance, as roll-call voting is often observed in reality, particularly in legislatures.

The equivalence result of Dekel and Piccione, though, crucially depends on the requirement of exogenous voting sequence. One can easily imagine, and perhaps design, voting procedures which allow voters more flexibility with respect to timing. This allows the possibility of learning the private signals by the later voters. In the context of private investment decision, Chamley and Gale [11] examine how social learning can affect the behaviour of late investors, and thus induces strategic delay for better information. In this chapter, we take a similar approach in a collective decision setting. We consider a multi-round voting procedure in which each voter can exercise her vote in any one of several successive rounds. After each round, voters receive information about the votes that have been cast in previous rounds. We show that if voters have identical preferences, there always exists a perfect Bayesian equilibrium which aggregates all the information into the decision efficiently. Moreover, if the model is symmetric, there is a *symmetric* equilibrium which achieves this end. The basic intuition is that each voter can convey information to others not only through her choice, but also the *timing* of that choice. Those with relatively strong information can vote early, while those with relatively weak information wait to see how things unfold. The following example clarifies how it works.

We construct a very simple example that may help in indicating how multi-round voting can improve efficiency of the outcome over simultaneous or Dekel and Piccione type sequential voting with an exogenous sequence. We consider a 3-voter binary election where the choice is between alternative (A) and status-quo (Q). Each voter can get one of the three signals $-1, 0$ or 1. The signal 0 is interpreted as a completely uninformative signal while the signal $1(-1)$ is a strong, though imperfect, signal in favour of $A(Q)$. The prior on A is such that the efficient outcome is the following:

$$d^*(\theta) = A \text{ iff } \theta \in \left\{ \begin{array}{c} (1,1,1),(1,1,0),(1,1,-1) \\ (1,1,0),(1,0,-1),(0,0,0) \end{array} \right\}$$

where θ is the realized signal vector and $d^*(\theta)$ is the collective decision. A wins if there are at least two votes in favour of A.

If voting is simultaneous, the voting strategy of player i is a function $s_i : \Theta_i \longrightarrow \{A, Q\}$ where $\Theta_i = \{1, 0, -1\}$ is the set of signals that player i can receive. Now, if in any symmetric equilibrium of the simultaneous voting game $d(0,0,0) = A$, then for this equilibrium strategy profile $d(0,0,-1) = A$. This outcome is clearly inefficient.

Let us consider a sequential voting procedure for the above example when the voting sequence is exogenous. Without loss of generality, we consider voter 1 as the first round voter. The voting strategy of voter 1 is still a function $s_1 : \Theta_1 \longrightarrow \{A, Q\}$, but for players 2 and 3, voting strategies depend on the history of the game. For later players who vote in rounds 2 and 3, the voting strategy is $s_i^t : H^t \times \Theta_i \longrightarrow \{A, Q\}$, where H^t is the set of all possible histories prior to period t. We will consider monotonic strategies. Voter 1 plays one of the following two strategies:

$$s_1^1 = \begin{cases} A & \text{if } \theta_1 = 1 \text{ or } 0 \\ Q & \text{if } \theta_1 = -1 \end{cases}$$

or

$$s_1^2 = \begin{cases} A & \text{if } \theta_1 = 1 \\ Q & \text{if } \theta_1 = 0 \text{ or } -1 \end{cases}.$$

Suppose voter 1 plays s_1^1. If $\theta_1 = 1$ is followed by $\theta_2 = 0$ and $\theta_3 = -1$, then for the outcome to be efficient at least one among voters 2 and 3 have to vote for A since $d^*(1, 0, -1) = A$. Now suppose $\theta_1 = 0$ is followed by $\theta_2 = 0$ and $\theta_3 = -1$. Voters 2 and 3 now face the same histories and private signals and therefore vote similarly. Hence, A still wins although $d^*(0, 0, -1) = Q$. A similar inefficiency can be shown to exist if voter 1 plays s_1^2 in equilibrium.

In the game with flexible timing, the existence of a symmetric efficient equilibrium is easy to show. Only two rounds of voting are needed. In round 1, each player has three actions to choose from – A, Q and W, where W stands for waiting in round 1. After round 1, the vote tallies in favour of A and Q become public knowledge and round 2 voters condition their vote on the history. Suppose (m_A, m_Q) denote the number of votes in favour of A and Q respectively after round 1. The following strategy now constitutes an efficient symmetric equilibrium:

$$s_i^1 = \begin{cases} A & \text{if } 1 \\ Q & \text{if } -1 \\ W & \text{if } 0 \end{cases}$$

and

$$s_i^2 = \begin{cases} A & \text{if } (1-1) \\ A & \text{if } (1-0) \\ A & \text{if } (0-0) \\ Q & \text{if } (0-1) \end{cases}$$

where s_i^2 is the action chosen by player i in round t. These strategies generate efficient outcomes for all possible signal realizations. After round 1, the private signals become common knowledge and the later voters vote to ensure the efficient outcome. Since voters have common preferences and information is aggregated efficiently, these strategies constitute an equilibrium.

What would happen if the preferences are diverse in a voting game with flexible timing? This will be a more difficult question to answer since even with pre-play communication the voters will have incentive to strategically withhold private information. In Section 3 of this chapter, we show that the multi-round voting mechanism may work in environments with diverse preferences as well. This mechanism achieves precisely what single round pre-play communication or a straw vote like opinion poll can achieve. But unfortunately we don't have a general relationship between the set of equilibria of the voting game with communication and the set of equilibria of the voting game with flexible timing. We conjecture that there would be a strong connection between the set of equilibria with unrestricted communication and that in the flexible timing game. But proving it is an agenda for further research.

When the size of the electorate is very large, the probability of a single player being pivotal is almost negligible and the notion of strategic voting does not have much appeal. Also communication among the electorate is almost impossible. On the other hand, in very small elections, it's hard to believe that players are anonymous and cannot talk prior to voting. This gives players access to unrestricted communication. So our model is mostly applicable in medium-sized elections where communication is costly. Possible examples can be election of faculty council or vote on some agenda by the residents of a medium sized city. We conclude that mechanisms that allow for a more flexible choice of timing could substantially improve the efficiency of collective decisions.

A common example of a sequential voting procedure is the roll-call voting often observed in legislative voting. However, in roll-call voting, voting sequence is exogenous. Hence, given the pivotal voter argument, the scope of the later voters to learn some extra information from the actions of the earlier voters is limited. We, on the other hand, allow voters to choose the timing of their votes thus allowing the earlier voters to convey their private information to the later participants.

In a different context, the papal election in Vatican proceeds in multiple rounds. In the papal election, there is no restriction in the number of candidates. Each voter from the College of Cardinals votes in the first round. If one candidate gets two-third majority in the first round, then he wins. If nobody gets the required majority in the first round, voting proceeds to the next round and the same procedure is repeated. The election mechanism used in the papal election, while not being the same as what we propose, has a similar flavour. After a round of voting with no clear winner, each voter updates his information using the voting history before the next round of ballot takes place.

We essentially ask a normative question in this chapter. We suggest an election mechanism where voters at each stage can update their information from the voting history of the preceding stages. Moreover, by strategically choosing their timing, they can now convey their private information to others if they so desire. We seek to examine the efficiency implications of such a mechanism and it turns out that a multiround voting mechanism does a pretty good job in aggregating disparate private information of the voters.

The chapter is organized in the following way. Section 2 describes the model and the results when preferences are aligned. Section 3 discusses an example with diverse preferences and shows the connection between the set of equilibria with communication and that with flexible timing. Section 4 concludes.

3.2 The Model with Common Interest

A group of n voters must choose from among two options—the status quo (Q) and the alternative (A). The state-of-the-world can also have two values, Q and A. All voters have common preferences captured by the utility function $u(d, \omega)$, where $d \in \{Q, A\}$ represents the voting outcome and ω is the state-of-the-world. We assume

$$u(d,\omega) = \begin{cases} -q & \text{if } d = A, \omega = Q \\ -(1-q) & \text{if } d = Q, \omega = A \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Voter i receives a private signal $\theta_i \in \Theta_i$. Let $\Theta = \Theta_1 \times \Theta_2 \times ... \times \Theta_n$, with θ being a typical element. We assume that the number of elements in Θ_i is finite. Let $f(\theta|\omega)$ denote the probability of realization of the signal vector θ if the true state is ω. The prior on the event $\omega = A$ is π.

The voting rules are as follows. The alternative A is selected if the final vote tally in favour of A is k or more. Otherwise, Q is selected. Voting proceeds in T successive rounds. Each voter has a *single*, indivisible vote, which she must cast in favour of either Q or A in any one of the T rounds. Thus in rounds $1, 2, ..., T-1$, a voter who has not previously cast her vote must choose from a set of three alternatives: $\{Q, A, W\}$, where W stands for "wait till the next round." In round T, the set of options is simply $\{Q, A\}$. Votes, once cast, cannot be changed subsequently. In general, let X_i^t denote the set of options available to voter i at date t. If i has exercised her vote before, we adopt the convention that $X_i^t = \{W\}$.

While each voter's signal is private information, the actions of all voters in each round become common knowledge before the next round[17]. Voter i's

[17]Since we are interested in environments in which voters cannot directly communicate with each other, the requirement that they are able to observe the actions of *all* other voters after each round may seem restrictive. However, in the symmetric case, all that is necessary for our results is that voters be able to observe some relevant summary statistic of other voters' actions, e.g. the *aggregate* number of votes cast in favour of either alternative after each round.

history up to round t, denoted by h_i^t, is the sequence of actions taken by i in rounds $1, 2, ..., t$. Let h^t be a vector of histories of all players, and H^t the set of all possible histories up to round t, with $H^0 = \Phi$. A (pure) strategy s_i for player i is then a collection of mappings $s_i^t : H^{t-1} \times \Theta_i \longrightarrow X_i^t$ for $t = 1, 2, ..., T$.

The above discussion describes a well defined Bayesian game. We focus on perfect Bayesian equilibria (PBE) of this game.

Before proceeding with characterizing equilibria, it is useful first to define a benchmark. Suppose that all players' signals were common knowledge. Then, voters will have a common posterior defined by Bayes' rule as follows:

$$\gamma(\theta) = \Pr[\omega = A|\theta] = \frac{\pi f(\theta|\omega = A)}{\pi f(\theta|\omega = A) + (1-\pi) f(\theta|\omega = Q)} \tag{3.2}$$

Let $d^*(\theta)$ denote the optimal choice, given the entire vector of signals. Clearly

$$d^*(\theta) = A \Longleftrightarrow \gamma(\theta) \geq q \tag{3.3}$$

For every profile of strategies s in the voting game, and for every realized signal vector θ, we can define $\widehat{d}(s, \theta) \in \{Q, A\}$ as the outcome induced by s.

Definition 3.1. *A strategy profile s satisfies* full information equivalence *if $\widehat{d}(s, \theta) = d^*(\theta)$ for all $\theta \in \Theta$.*

Finally, denote by m_i, the number of elements in Θ_i, and let $\overline{m} = \max_{i \in \mathcal{I}} m_i$. We are now ready to state the main result of the chapter.

Theorem 3.1. *Suppose $T > \overline{m}$. Then there exists a pure strategy perfect Bayesian equilibrium $\widehat{s}$ of the voting game which satisfies full information* *equivalence.*

Proof. To prove this result, it is enough to show that there exists a strategy profile which achieves full information equivalence. A useful property of common interest games is that the profile of strategies which maximizes the common payoff function is immune to individual deviations, i.e. constitutes a PBE[18]. Hence, if full information equivalence is feasible, it is achievable in equilibrium.

Demonstrating the existence of such a strategy profile is extremely simple. Consider the strategy profile $\widehat{s}$ as follows. Divide the voters into three subsets: $\mathcal{I}_A = \{1, 2, ..., k-1\}$, $\mathcal{I}_Q = \{k, ..., n-1\}$ and $\mathcal{I}_p = \{n\}$. Let $n_A(h^t)$ and $n_Q(h^t)$ be the number of votes cast in favour of A and Q respectively under history h^t. Also, let $\Theta_i = \{1, 2, ..., m_i\}$. For $i \in \mathcal{I}_A$

$$\widehat{s}_i^t(h^t, \theta_i) = \begin{cases} A & \text{if } t = \theta_i \\ W & \text{if } t \leq \overline{m} \text{ and } t \neq \theta_i \\ Q & \text{if } t = \overline{m} + 1, n_A(h^t) < k - 1, n_Q(h^t) < n - k \\ A & \text{if } t = \overline{m} + 1, n_A(h^t) = k - 1, n_Q(h^t) < n - k \\ Q & \text{if } t = \overline{m} + 1, n_A(h^t) < k - 1, n_Q(h^t) = n - k \\ Q & \text{otherwise} \end{cases} \tag{3.4}$$

[18]See McLennan [41] and Chakraborty and Ghosh [10].

Similarly, for $i \in \mathcal{I}_Q$

$$\widehat{s}_i^t(h^t, \theta_i) = \begin{cases} Q & \text{if } t = \theta_i \\ W & \text{if } t \leq \overline{m} \text{ and } t \neq \theta_i \\ Q & \text{if } t = \overline{m}+1, n_A(h^t) < k-1, n_Q(h^t) < n-k \\ A & \text{if } t = \overline{m}+1, n_A(h^t) = k-1, n_Q(h^t) < n-k \\ Q & \text{if } t = \overline{m}+1, n_A(h^t) < k-1, n_Q(h^t) = n-k \\ Q & \text{otherwise} \end{cases} \tag{3.5}$$

Finally for $i \in \mathcal{I}_p$

$$\widehat{s}_i^t(h^t, \theta_i) = \begin{cases} W & \text{if } t \leq \overline{m} \\ A & \text{if } t = \overline{m}+1, n_A(h^t) = k-1, n_Q(h^t) = n-k \\ & \quad \text{and } \Pr\left[\omega = A | \widehat{s}_{-i}, h^t, \theta_i\right] \geq q \\ Q & \text{if } t = \overline{m}+1, n_A(h^t) = k-1, n_Q(h^t) = n-k \\ & \quad \text{and } \Pr\left[\omega = A | \widehat{s}_{-i}, h^t, \theta_i\right] < q \\ A & \text{if } t = \overline{m}+1, n_A(h^t) < k-1, n_Q(h^t) < n-k \\ A & \text{if } t = \overline{m}+1, n_A(h^t) = k-1, n_Q(h^t) < n-k \\ Q & \text{if } t = \overline{m}+1, n_A(h^t) < k-1, n_Q(h^t) = n-k \\ A & \text{otherwise} \end{cases} \tag{3.6}$$

Since the first $k-1$ voters vote for A, and the next $n-k$ voters vote for Q regardless of their signals, the last voter is always pivotal along the equilibrium path of play. Although the first $n-1$ votes are independent of the signals, their *timing* bears a one-to-one correspondence with the signals. Hence, voter n can infer the signal realizations of all preceding voters by date $t = \overline{m}+1$. Since the last and pivotal voter votes optimally with respect to her equilibrium information, the outcome satisfies full information equivalence. Since the equilibrium payoff is maximum that a voter can get in the above game, these strategies are immune from individual deviations under any set of off–the-equilibrium beliefs.

We adopt a particular set of off-the-equilibrium beliefs and strategies, but many other specifications would work as well. In words, it can be described as follows. It is common knowledge that there has been a deviation if one of the non-pivotal voters hasn't voted by date $\overline{m}$. Assume that the votes of others will be interpreted as if they have played in accordance with their equilibrium strategies, and the votes of those who have waited will be seen as uninformative. Further, at every date after it becomes clear that a deviation has taken place, everyone must simultaneously vote in favor of Q, unless there is an unanimity required among the remaining voters to adopt Q, in which case they simultaneously and unanimously vote for A. If there is a single voter left, she votes for whichever outcome she sees as optimal, given her beliefs at that point. Since, as long as there are two or more voters left, no voter is pivotal under this construction, every voter plays a best response. ∎

Since we are interested in environments in which voters cannot communicate and exchange information directly, the relevance of the above formulation may be called into question for two reasons. First, it requires voters to have individual specific information about each voter's actions in the previous rounds. This may seem too strong a requirement. It may be more reasonable to assume that voters get information about *aggregate* vote tallies from the past rounds, not *who* voted and who abstained. Second, it is clear from the strategy constructed above that the partition of the set of voters, including the identity of the last and pivotal voter, is arbitrary—any permutation would work just as well. In other words, even if voters are ex ante identical, they must play asymmetric strategies based on a pre-specified partition that could be selected in many different ways. This requires solving a complex coordination problem which seems both unrealistic and incompatible with the assumption of lack of communication. Nevertheless, our next result shows that in a symmetric model, subject to a mild condition on the joint distribution over signals being satisfied, it is possible to construct a *symmetric* strategy profile which successfully aggregates all the information.

Henceforth, suppose each player could receive one among m signals, so that $\Theta_i = \{1, 2, ..., m\}$. The probability distribution $f(\theta|\omega)$ is *symmetric* if the value of the function is the same for any permutation of the components of θ. $f(\theta_i, \theta_{-i}|\omega)$ is *monotonic* if $\frac{f(\theta_i,\theta_{-i}|A)}{f(\theta_i,\theta_{-i}|Q)}$ is monotonic in θ_i for all $\theta_{-i} \in \Theta_{-i}$. Without loss of generality, we assume that $\frac{f(\theta_i,\theta_{-i}|A)}{f(\theta_i,\theta_{-i}|Q)}$ is increasing in θ_i.

Theorem 3.2. *Assume $\Theta_i = \{1, 2, ..., m\}$ for all i, and $f(\theta|\omega)$ is symmetric and monotonic. If $T \geq m$, there is a symmetric PBE in pure strategies which satisfies full information equivalence.*

Before proceeding with the proof, a few definitions and preliminary results need to be established. Consider any subset of l voters and a subset $\Theta' \subseteq \Theta$ of signals. Let Θ'_l denote the product space of signals received by l members, each signal being drawn from the subset Θ'. Let $\overline{\theta}$ denote the highest signal in Θ' and $\underline{\theta}$ the lowest. Let $\overline{\Theta}'(r)$ denote the subset of Θ' such that each member consists of at least r highest (i.e. $\overline{\theta}$) signals and the remaining arbitrary. Similarly, let $\underline{\Theta}'(r)$ denote the subset of Θ' such that each member consists of at least r lowest (i.e. $\underline{\theta}$) signals and the remaining arbitrary. Define

$$\overline{\gamma}(r; l, \Theta') = \min_{\theta \in \overline{\Theta}'(r)} \gamma(\theta) \tag{3.7}$$

$$\underline{\gamma}(r; l, \Theta') = \max_{\theta \in \underline{\Theta}'(r)} \gamma(\theta) \tag{3.8}$$

Let $\theta_{s,l-s} \in \Theta'_l$ denote a signal vector that consists of s number of $\overline{\theta}$ signals and $l - s$ number of $\underline{\theta}$ signals. Due to monotonicity, it follows that

$$\overline{\gamma}(r; l, \Theta') = \underline{\gamma}(l - r; l, \Theta') = \gamma(\theta_{r,l-r})$$

and further, $\overline{\gamma}(r; l, \Theta')$ is increasing in r, while $\underline{\gamma}(r; l, \Theta')$ is decreasing in r.

Proof. (of Theorem) We can now complete the the proof of the theorem. Consider the first round. Either $\bar{\gamma}(k; n, \Theta) \geq q$ or $\underline{\gamma}(n - k + 1; n, \Theta) < q$. Suppose $\bar{\gamma}(k; n, \Theta) \geq q$. Then, let all the voters who received a signal m vote for A in this round, and the remaining choose W. If the outcome is decided by these votes straight away, that means at least k voters must have received the signal m. In that case, the updated posterior on the state being A, conditioning on the entire vector of signals, is bounded below by $\bar{\gamma}(k; n, \Theta) \geq q$. Then the decision is efficient, regardless of what the other signals are. On the other hand, if the number of votes cast is less than k, the game proceeds to the next round. If $\bar{\gamma}(k; n, \Theta) < q$, then the strategy is to vote for Q in the first round if a voter has received the signal 1, and to wait otherwise.

Suppose, after application of these first round strategies, the result of the election is still undecided, so that the game proceeds to round 2. Exactly similar strategies can now be constructed recursively. Notice that since the number of voters who voted in the previous round, as well as their signals (through knowledge of their strategies) is commonly known, all remaining voters can update their beliefs over the states accordingly. The subgame beginning with round 2 is exactly similar to the initial game, only with a reduced set of players (after eliminating those who have already voted), a reduced set of possible signals for each, an updated prior as well as probability distribution over signals, and a new threshold of votes necessary for each alternative to be chosen. The same exercise as in round 1 can be applied to this subgame to construct round 2 strategies after every possible history. This recursive method can last at most m steps before it yields an outcome, and by construction, the outcome is efficient whenever the process stops.

More formally, after any history h_t, let $n(h_t)$ be the number of voters left, $\Theta(h_t)$ the reduced set of possible signals, and $k((h_t)$ the number of remaining votes needed for A to be chosen. Either $\bar{\gamma}(k(h_t); n(h_t), \Theta(h_t)) \geq q$ or $\underline{\gamma}(k(h_t); n(h_t), \Theta(h_t)) < q$. Let $\bar{\theta}(h_t) = \max \Theta(h_t)$ and $\underline{\theta}(h_t) = \min \Theta(h_t)$. Now if $\bar{\gamma}(k(h_t); n(h_t), \Theta(h_t)) \geq q$, let all the voters who receive signal $\bar{\theta}(h_t)$ vote for A, and all the others who didn't cast their votes in any of the previous rounds choose W. If $\underline{\gamma}(k(h_t); n(h_t), \Theta(h_t)) < q$, voters who receive $\underline{\theta}(h_t)$ vote for Q and wait otherwise. This strategy ensures that after any history h_t, if the outcome is decided, then it is efficient. Since, as a result of this strategy profile, the best outcome is reached, this is a perfect Bayesian equilibrium. ∎

The following example clarifies how the process described above works.

We consider a simple voting game where $n = 3$, $\Theta_i = \{1, 2, 3\}$ and outcome is decided by majority rule in favour of A, i.e. $k = 2$.

Case I: $\gamma(3, 3, 1) \geq q$

Anyone with signal 3 will vote for A in round 1. If the outcome is decided in the first round then it is efficient independent of the signals of the later voters, since $\gamma(3, 3, 3) > \gamma(3, 3, 2) > \gamma(3, 3, 1) \geq q$ by monotonicity. Otherwise, there are two possible histories at the end of round one: (i) one vote for A from which the later voters can perfectly infer that there is one 3 signal; (ii) zero vote for A.

In case of (ii), there are now only four possible signal realizations[19]:

$$\{2, 2, 2\}, \{2, 2, 1\}, \{2, 1, 1\}, \{1, 1, 1\}.$$

If $\gamma(2, 2, 1) \geq q$, all with signal 2 vote for A in round 2. Again if the outcome is determined, then it is efficient. If not, then in the last round the remaining voters know the exact realization of the signal vector and vote accordingly. If $\gamma(2, 2, 1) < q$, all with signal 1 vote for Q. The same argument as above shows that outcome is efficient.

In case of (i), there are three possible realizations of signal vector:

$$\{3, 2, 2\}, \{3, 2, 1\}, \{3, 1, 1\}.$$

Now the number of votes required for A to be chosen is one. If $\gamma(3, 2, 1) \geq q$, voters with signal 2 vote for A in round 2. If $\gamma(3, 2, 1) < q$, voters with signal 1 vote for Q in this round. The outcome is efficient by similar reason as in

[19]Because of symmetry, the permutations of signals can be ignored.

(ii).

Case II: $\gamma(3,3,1) < q$

Anybody with signal 1 votes for Q in round 1. Outcome is efficient if decided in period 1. If it is still undecided, similar strategy as in Case I will generate efficient outcome.

3.3 Diverse Preference

In this section, we will try to explore the effectiveness of our mechanism in a situation where people's preferences may vary. Even when preferences vary, each player has incentive to make the best informed choice and hence to share their private information. On the other hand, since interests can be conflicting, revealing one's private information completely may prove too costly. These two opposing incentive will lead to some revelation of private information in equilibrium if communication between the players are permitted. Doraszelski, Gerardi and Squintani [19] (henceforth DGS) have studied this situation in a two voter binary signal model. In their model, prior to the voting stage, the players may communicate using a straw vote to send messages to other players. This takes the form of an opinion poll prior to voting. In our case, we do not allow people to communicate among each other. In stead, we use the indirect mechanism of multi-round voting where voters get to choose the timing of their votes. We will show, using the same model as DGS, that the best equilibria in their model with communication can be easily replicated using our mechanism. As we have already shown in the earlier sections of this chapter, when interests are completely aligned, multi-round voting achieves what unrestricted communication can do as an equilibrium outcome. The purpose of this section is to show that the usefulness of our mechanism is not restricted to the special case of common interest. This mechanism can substitute for pre-play communication under more general circumstances.

We use the same model as in Section 3.2 except that we now have only two voters and voters may differ in their relative concerns for two types of errors that may occur as a consequence of a wrong decision. A player of type q has the same utility function as in equation [3.1]. Types are private information, but it is common knowledge that each type is an i.i.d draw from a distribution F with domain $(0,1)$. F is continuous and strictly increasing. We assume that players are not extreme types where they completely ignore one types of error.

Each player i receives a private signal $s_i \in \{0,1\}$ which is related to the state of the world: $\Pr(s_i = 1|A) = \Pr(s_i = 0|Q) = p \in \left(\frac{1}{2}, 1\right)$. The signals are conditionally independent. Voting proceeds in two rounds[20]. After the first round, voting history is revealed to anyone who decides to vote in round two. Finally, conviction requires unanimity[21].

We begin by discussing the equilibria of the game with communication. Both the players can exchange messages prior to voting. There always exist some equilibria where players ignore the messages completely and the game proceeds just like a pure voting game. These equilibria are called babbling equilibria. We ignore this class of equilibria. We concentrate on the set

[20]Since in this section we are restricting ourselves to a binary signal space, we need just two rounds of voting.

[21]The restriction to a particular voting rule is to ensure that we are using the same model as DGS.

of Perfect Bayesian equilibria in cutoff strategies that do not admit weakly dominated strategies. It can be shown that all the Perfect Bayesian equilibria of the communication game are outcome equivalent to the set of perfect Bayesian equilibria in cut off strategies[22].

Not all cutoffs are necessarily identified by sequential rationality or weak dominance. In the case an action by any of the player affects the final decision, the cutoffs are uniquely identified across signals by a family of functions

$$k(q) = \frac{(1-p)^2 q}{(1-p)^2 q + p^2(1-q)}$$

defined on $(0,1)$. Note that $k'(q) > 0$ and $k(q) < q$ for all q. The equilibria which admit cutoff strategies thus related across signals are called *robust.* Moreover, in equilibrium, each player conditions her vote on her preference parameter q, the signal she receives s, the message she sends m and the message she receives M. These equilibria are called *responsive.* Finally, we consider only the set of symmetric PBEs.

The equilibrium profiles are identified by the cutoffs q_s and q_{smM} with $s \in \{0,1\}, m \in \{0,1\}$, and $M \in \{0,1\}$. This means that a player of type q sends message $m = 1 (m = 0)$ after observing signal s if $q < q_s$ $(q > q_s)$, and that a player of type q votes $v = A$ $(v = Q)$ after observing signal s, sending message m, and receiving message M if $q < q_{smM}$ $(q > q_{smM})$.

Proposition 3.1. (DGS) *There exist three classes of responsive robust cutoff equilibria:*

Class 0: $q_s < q_{s10} < q_{s00} < q_{s11} = q_{s01}$ *for* $s \in \{0,1\}$;

[22] See DGS [19] for a detailed discussion.

Class 1: $q_{s00} < q_{s10} < q_s < q_{s11} < q_{s01}$ *for* $s \in \{i, g\}$*;*

Class 2: $q_{s00} < q_{s10} = q_s < q_{s11} = q_{s01}$ *for* $s \in \{i, g\}$*.*

There does not exist a responsive robust cutoff equilibrium in any other configuration.

For any class, the smallest cutoff in $s = 0$ *is larger than the largest cutoff in* $s = 1$ *and the two sets of cutoffs are separated by* $\frac{1}{2}$*. Moreover, the equilibria of class 0 are outcome-equivalent to the equilibria of class 1.*

This proposition is directly taken from DGS [19] with some minor changes to make it compatible with the notations of our model. The proof is omitted and can be found in DGS [19]. It can also be shown that class 0 and class 1 equilibria are ex-ante Pareto superior to class 2 equilibria. We will show that any class 0 equilibrium is outcome equivalent to a Perfect Bayesian equilibrium of the multi-round voting game.

We now turn to the equilibria of class 0. Figure 3.1 illustrates the equilibrium behaviour. For each interval, the first column reports the path of play of the type when she has received the signal $s = 1$ from nature, and the second column, the path after the signal $s = 0$. The first row identifies the message sent. The second row refers to the vote. When a type conditions her vote on her opponents message, we first present the vote after receiving message $M = 1$ and second the vote after $M = 0$.

The cutoffs described in the above proposition for class 0 equilibria are defined by the following set of equations [19]:

$$q_{sm1} = \frac{1}{1 + R\frac{[RF(q_s)+F(k(q_s))]}{[F(q_s)+RF(k(q_s))]}} \tag{3.9}$$

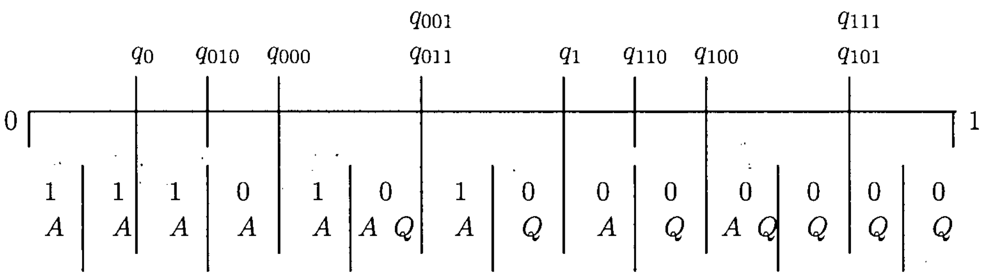


Figure 3.1: Equilibrium path of play for class 0 equilibria in DGS

and

$$q_{sm0} = \frac{1}{1 + R\frac{[R(F(q_{sm0})-F(q_s))+(F(k(q_{sm0}))-F(k(q_s)))]}{[(F(q_{sm0})-F(q_s))+R(F(k(q_{sm0}))-F(k(q_s)))]}} \tag{3.10}$$

and, finally,

$$q_s = \frac{1}{1 + R\frac{R(F(q_{sm1})-F(q_{sm0}))+(F(k(q_{sm1}))-F(k(q_{sm0})))}{(F(q_{sm1})-F(q_{sm0}))+R(F(k(q_{sm1}))-F(k(q_{sm0})))}} \tag{3.11}$$

for $s \in \{0,1\}$. It should be noted that other configurations of equilibrium cutoffs are possible. We select this class of equilibria since this is ex-ante Pareto efficient as shown by DGS.

Coming back to the analysis of multi-round game, we will try to find a symmetric Perfect Bayesian Equilibrium in cut-off strategies. Since $d = A$

requires unanimity, any voter who votes for Q can do so in round two. A single voter can decide the outcome in favour of Q. Hence, a Q voter loses nothing by waiting in round one.

We consider an equilibrium in the following symmetric cutoff structure. A player with signal $s \in \{0, 1\}$ will vote A in round one if $q \leq \hat{q}_s$. Otherwise, the voter waits. A waiting voter can face two possible histories in round two: (i) h^0 - 0 votes cast in round one and (ii) h^1 - 1 vote cast in round one. Confronted with history $h^n, n \in \{0, 1\}$, a voter with signal s will vote A in round two if $q \leq \hat{q}_{sn}$.

For the above cutoff strategies to be an equilibrium, $\hat{q}_s < \hat{q}_{sn}$ must hold for $n \in \{0, 1\}$ and $s \in \{0, 1\}$. In round two, voter i will vote A if and only if $\Pr\left[A|piv, h^n, s_i\right] \geq q_i$. Hence, $\hat{q}_{sn} = \Pr\left[A|piv, h^n, s\right]$.

Lemma 3.1. *Given any history h^n, the cutoffs are related across signals by the function:*

$$\hat{q}_{0n} = k(\hat{q}_{1n}) = \frac{(1-p)^2 \hat{q}_{1n}}{(1-p)^2 \hat{q}_{1n} + p^2(1-\hat{q}_{1n})} \tag{3.12}$$

for $n = 0, 1$. Moreover, $\hat{q}_{0n} < \hat{q}_{1n}$ for $n \in \{0, 1\}$.

Proof. We can write

$$\begin{aligned}
\hat{q}_{1n} &= \Pr\left[A|piv, h^n, 1\right] \\
&= \frac{\Pr\left[piv, h^n, 1|A\right].\Pr[A]}{\Pr\left[piv, h^n, 1|A\right].\Pr[A] + \Pr\left[piv, h^n, 1|Q\right].\Pr[Q]} \\
&= \frac{1}{1 + \frac{\Pr[piv, h^n, 1|Q].\Pr[Q]}{\Pr[piv, h^n, 1|A].\Pr[A]}} \\
&= \frac{1}{1 + \frac{\Pr[1|Q]}{\Pr[1|A]}\cdot\frac{\Pr[piv, h^n|Q]}{\Pr[piv, h^n|A]}} \\
&= \frac{1}{1 + \frac{1-p}{p}\cdot\frac{\Pr[piv, h^n|Q]}{\Pr[piv, h^n|A]}}.
\end{aligned}$$

The above holds since signals are conditionally independent and uncorrelated to q and also $\Pr\left[A\right] = \Pr\left[Q\right] = \frac{1}{2}$. This can be simplified to

$$\frac{\Pr\left[piv, h^n|Q\right]}{\Pr\left[piv, h^n|A\right]} = \frac{p}{1-p}.\frac{1-\hat{q}_{1n}}{\hat{q}_{1n}}$$

Similarly,

$$\begin{aligned}
\hat{q}_{0n} &= \Pr\left[G|piv, h^n, 0\right] \\
&= \frac{1}{1 + \frac{p}{1-p}\cdot\frac{\Pr[piv, h^n|Q]}{\Pr[piv, h^n|A]}} \\
&= \frac{(1-p)^2\hat{q}_{1n}}{(1-p)^2\hat{q}_{1n} + p^2(1-\hat{q}_{1n})} \\
&= k\left(\hat{q}_{1n}\right) < \hat{q}_{1n}.
\end{aligned}$$

for any $\hat{q}_{1n} \in (0, 1)$. ∎

After history h^1, the waiting player is pivotal with probability one. Hence,

$$\begin{aligned}
\hat{q}_{11} &= \frac{1}{1+\frac{1-p}{p}\cdot\frac{\Pr[piv,h^1|Q]}{\Pr[piv,h^1|A]}} \\
&= \frac{1}{1+\frac{1-p}{p}\cdot\frac{\Pr[h^1|Q]}{\Pr[h^1|A]}} \\
&= \frac{1}{1+\frac{1-p}{p}\cdot\frac{(1-p)F(\hat{q}_1)+pF(\hat{q}_0)}{pF(\hat{q}_1)+(1-p)F(\hat{q}_0)}} \\
&= \frac{1}{1+R.\frac{RF(\hat{q}_1)+F(\hat{q}_0)}{F(\hat{q}_1)+RF(\hat{q}_0)}}
\end{aligned} \tag{3.13}$$

where $R = \frac{1-p}{p} < 1$. By Lemma 3.1, $\hat{q}_{01} = k(\hat{q}_{11})$.

After history h^0, the probability of being pivotal is less than one. An waiting player i can observe history h^0, if either the other player's type $q^j > \hat{q}_0$ coupled with $s_j = 0$ or $q^j > \hat{q}_1$ coupled with $s_j = 1$. In case of the first event player i is pivotal under the symmetric cutoff strategies only when $q^j \in (\hat{q}_0, \hat{q}_{00}]$. If $s_j = 1$, player i is pivotal when $q^j \in (\hat{q}_1, \hat{q}_{10}]$. We can now write

$$\Pr\left[piv, h^0|A\right] = p\left[F(\hat{q}_{10}) - F(\hat{q}_1)\right] + (1-p)\left[F(\hat{q}_{00}) - F(\hat{q}_0)\right]$$

and

$$\Pr\left[piv, h^0|Q\right] = (1-p)\left[F(\hat{q}_{10}) - F(\hat{q}_1)\right] + p\left[F(\hat{q}_{00}) - F(\hat{q}_0)\right].$$

Therefore,

$$\begin{aligned}
\hat{q}_{10} &= \frac{1}{1+\frac{1-p}{p}\cdot\frac{\Pr[piv,h^0|Q]}{\Pr[piv,h^0|A]}} \\
&= \frac{1}{1+\frac{1-p}{p}\cdot\frac{(1-p)[F(\hat{q}_{10})-F(\hat{q}_1)]+p[F(\hat{q}_{00})-F(\hat{q}_0)]}{p[F(\hat{q}_{10})-F(\hat{q}_1)]+(1-p)[F(\hat{q}_{00})-F(\hat{q}_0)]}} \\
&= \frac{1}{1+R.\frac{R[F(\hat{q}_{10})-F(\hat{q}_1)]+[F(\hat{q}_{00})-F(\hat{q}_0)]}{[F(\hat{q}_{10})-F(\hat{q}_1)]+R[F(\hat{q}_{00})-F(\hat{q}_0)]}}. \qquad (3.14)
\end{aligned}$$

Again by Lemma 3.1, $\hat{q}_{00} = k(\hat{q}_{10})$.

We still have to determine the round 1 cutoffs. Since in equilibrium $\hat{q}_{sn} > \hat{q}_s$ for $n = 0, 1$, the marginal type's choice is between voting A in round 1 and waiting to vote A in round 2. Hence, $\hat{q}_s$ is determined from the following equality:

$$EU(A \text{ in round } 1|s) = EU(A \text{ in round } 2|s)$$

Suppose player i decides to vote A in round 1. If the other player plays the symmetric cutoff strategies, then A will be the voting outcome in either of the following two cases:

(i) Player j receives an 0 signal and $q^j \leq \hat{q}_{01}$. If $q^j \leq \hat{q}_0$, then A will be decided in first round. If $q^j > \hat{q}_0$, player j waits in round 1; but since player j will be confronted with h^1, he will vote A in round 2 if $q^j \leq \hat{q}_{01}$.

(ii) Player j receives 1 and $q^j \leq \hat{q}_{11}$.

Hence, for a type q voter,

$$\begin{aligned}
& EU(A \text{ in round } 1|s) \\
= \; & \Pr(0|s)\left[F(\hat{q}_{01})V(A,s,0) - (1-F(\hat{q}_{01}))V(Q,s,0)\right] \\
+ \; & \Pr(1|s)\left[F(\hat{q}_{11})V(A,s,1) - (1-F(\hat{q}_{11}))V(Q,s,1)\right] \\
= \; & \Pr(0|s)\left[-F(\hat{q}_{01})(1-\gamma(s,0))q - (1-F(\hat{q}_{01}))\gamma(s,0)(1-q)\right] \\
+ \; & \Pr(1|s)\left[-F(\hat{q}_{11})(1-\gamma(s,1))q - (1-F(\hat{q}_{11}))\gamma(s,1)(1-q)\right]
\end{aligned}$$

where $V(d, s_i, s_j)$ is the expected utility from decision d when the signals are s_i, s_j and $\gamma(s_i, s_j) = \Pr[G|s_i, s_j]$. Similarly,

$$\begin{aligned}
& EU(A \text{ in round } 2|s) \\
= \; & \Pr(0|s)\left[-F(\hat{q}_{00})(1-\gamma(s,0))q - (1-F(\hat{q}_{00}))\gamma(s,0)(1-q)\right] \\
+ \; & \Pr(1|s)\left[-F(\hat{q}_{10})\gamma(s,1)q - (1-F(\hat{q}_{10}))(1-\gamma(s,1))(1-q)\right].
\end{aligned}$$

Now the cutoff $\hat{q}_s$ can be determined from the following equation:

$$EU(A \text{ in round } 1|s) = EU(A \text{ in round } 2|s)$$

$$\Rightarrow \Pr(0|s)\left(F(\hat{q}_{01}) - F(\hat{q}_{00})\right)\left[\hat{q}_s - \gamma(s,0)\right] = \Pr(1|s)\left(F(\hat{q}_{11}) - F(\hat{q}_{10})\right)\left[\gamma(s,1) - \hat{q}_s\right]$$

$$\Rightarrow \hat{q}_s = \frac{\Pr(1|s)\left(F(\hat{q}_{11}) - F(\hat{q}_{10})\right)\gamma(s,1) + \Pr(0|s)\left(F(\hat{q}_{01}) - F(\hat{q}_{00})\right)\gamma(s,0)}{\Pr(1|s)\left(F(\hat{q}_{11}) - F(\hat{q}_{10})\right) + \Pr(0|s)\left(F(\hat{q}_{01}) - F(\hat{q}_{00})\right)}.$$

Since $\Pr[A] = \Pr[Q] = \frac{1}{2}$, $\gamma(1,1) = \frac{p^2}{p^2+(1-p)^2}$, $\gamma(1,0) = \gamma(0,1) = \frac{1}{2}$, $\gamma(0,0) =$

$\frac{(1-p)^2}{p^2+(1-p)^2}$ and $\Pr(1|1) = \Pr(0|0) = p^2+(1-p)^2$, $\Pr(0|1) = \Pr(1|0) = 2p(1-p)$. Hence, simplifying we get

$$\hat{q}_1 = \frac{1}{1 + R.\frac{R(F(\hat{q}_{11})-F(\hat{q}_{10}))+(F(\hat{q}_{01})-F(\hat{q}_{00}))}{(F(\hat{q}_{11})-F(\hat{q}_{10}))+R(F(\hat{q}_{01})-F(\hat{q}_{00}))}} \tag{3.15}$$

and

$$\hat{q}_0 = k(\hat{q}_1).$$

Equations [3.13], [3.14] and [3.15] along with $\hat{q}_{0n} = k(\hat{q}_{1n})$ and $\hat{q}_0 = k(\hat{q}_1)$ now determine the equilibrium cutoffs. Notice that $k(x) < \frac{1}{2}$ for all $x > \frac{1}{2}$. Also, assuming that $\hat{q}_1$, $\hat{q}_{10}$ and $\hat{q}_{11}$ exist, it is easy to examine that all are greater than $\frac{1}{2}$. Hence, the highest cutoff corresponding to signal 0 is lower than the lowest cutoff corresponding to signal 1 and these two are separated by $\frac{1}{2}$.

From equation [3.14], we can express $\hat{q}_{10} = \phi(\hat{q}_{10}, \hat{q}_1)$. The function $\phi(., \hat{q}_1)$ is well-defined over the domain $(\hat{q}_1, 1]$ and continuous in $\hat{q}_{10}$. Since $\phi(., \hat{q}_1) > \hat{q}_1$ for all $\hat{q}_{10} \in (\hat{q}_1, 1]$, $\lim_{x\to\hat{q}_1} \phi(x, \hat{q}_1) > \hat{q}_1$ and $\lim_{x\to 1} \phi(x, \hat{q}_1) < 1$. We can now conclude that given any $\hat{q}_1 \in (0,1)$, $\hat{q}_{10}$ exists and greater than $\hat{q}_1$. We write $\hat{q}_{10} = h_0(\hat{q}_1)$. Similarly, from 3.13, we can express $\hat{q}_{11} = h_1(\hat{q}_1)$. Notice that in finding $\hat{q}_{10}$ and $\hat{q}_{11}$, we have made use of the relation $\hat{q}_0 = k(\hat{q}_1)$.

We still have to ensure the existence of $\hat{q}_1$. From equation [3.15], we express $\hat{q}_1 = \psi(\hat{q}_1)$. The function $\psi(.)$ is continuous and $\lim_{x\to 0} \psi(.) > 0$ and $\lim_{x\to 1} \psi(.) < 1$. An application of Brouwer's Fixed Point Theorem now ensures the existence of $\hat{q}_1$.

The equilibrium path of play for the multi-round voting game is characterized in Figure 3.2. The first column in each interval shows the player's

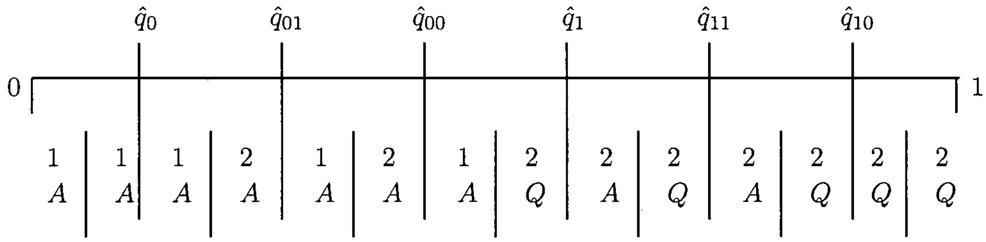


Figure 3.2: Equilibrium path of play in multi-round voting game

choice of timing and vote if the signal received is $s = 1$ and the second column shows the same for $s = 1$. The first row shows the period at which the voter votes and the second row represents the preferred alternative. We describe the equilibrium in the multi-round game in the following proposition.

Proposition 3.2. *The set of Perfect Bayesian equilibrium that admits a cutoff configuration in the multi-round voting game has the following cutoff structure:*

$$\hat{q}_s < \hat{q}_{s0} < \hat{q}_{s1}$$

A player with signal s votes for A in round 1 if her type $q \leq \hat{q}_s$. *If* $q > \hat{q}_s$, *the player waits and votes for A after observing history* h^n *iff* $q \leq \hat{q}_{sn}$.

We can now easily compare between the set of equilibria in the game with communication and the multi-round voting game. Since the same set of equations determine the relevant cutoffs for both games, the voting cutoffs are same. Hence the set of symmetric cut-off equilibria of the multi-round game is outcome equivalent to the set of efficient responsive robust cutoff equilibria in the game with communication.

This example illustrates that efficiency of flexibility of timing of votes in a voting game is not limited to a situation of aligned preferences. Our conjecture is that endogenizing the timing of votes can substitute for direct communication in aggregation of private information in a collective choice problem. Unfortunately, we don't have the general result yet. The link between the set of equilibria in these two types of games is not obvious. In our mechanism, to communicate her private information to others a player has to commit to a particular choice. In case of pre-play communication, a player can send a message and then deviate from her message in the voting stage. That is not feasible in the multi-round voting game.

3.4 Concluding Remarks

We show that in a common interest election, full information aggregation can be achieved if flexibility in timing of votes is allowed. With symmetric and monotonic signals, this can be achieved in a symmetric Perfect Bayesian Equilibrium. Moreover, the result does not depend on the degree and nature of correlation between private signals. The assumption of conditionally independent signals is an almost common feature of the literature of strategic voting, because otherwise the analysis becomes highly complicated. Our result is more general.

What would happen if the preferences are diverse in a voting game with flexible timing? This will be a more difficult question to answer since even with communication the voters will have incentive to strategically withhold private information. With diverse preferences, if the preferences are sufficiently close to induce full revelation of private information in voting with communication, the same can be achieved with the endogenous timing of votes. In the more general case, we conjecture that there would be strong connection between the set of equilibria in the game with communication and that in the flexible timing game. In Section 3 of this chapter, we illustrate that in the two person binary signal example. But proving it in a more general environment is an agenda for future research.

How this mechanism can be implemented when the number of voters are large enough to permit anonymity? Many elections are now held online with electronic voting. With advancement of Internet technology, it is now feasible to offer continuous real time updates of voting history. We conclude that if the objective of the election is information aggregation, this mechanism works well.

4. Committee Design with Endogenous Information

4.1 Introduction

A prime task of any organization is acquisition and processing of information relevant to important decisions. Informed decisions reduce the chance of errors; guesswork increases inefficiency. *Design* of committees within an organization for the purpose of information processing and collection is therefore very important.

The Condorcet Jury Theorem of the first kind (see McLean and Hewitt [40]) states that a majority of a group is more likely than a single individual to choose the better of the two alternatives. With exogenous information, the result is trivially true if perfect information aggregation tools are available. Since a group inherently possesses more information than a single individual, whenever the problem of information aggregation can be overcome the group can do no worse. Even when information must be aggregated through (possibly) imperfect mechanisms such as voting, the Condorcet result is valid, as illustrated in the literature on strategic voting (see Feddersen and Pessendorfer [21], Miller [43], or McLennan [41] for a sample). Chakraborty and Ghosh

[10] have provided a general result to this effect. In relation to committee design, these results show that with exogenous information a larger committee does better than a smaller committee under fairly general conditions.

What happens if information is endogenous? Since information is a public good, with endogenous information collection there could arise a *free-rider effect.* When several members are entrusted with the task of information collection, each may have an incentive to save private cost by collecting less information, free riding on the information of others. However, in some cases, there is another effect that shapes the incentive for information collection at the individual level. Since, in some situations, an individual's information may not be valuable in itself, but becomes valuable only in conjunction with information provided by others, individual pieces of information are often complementary to each other. This *information complementarity effect,* in some cases, bolsters the incentive to collect information in a larger committee and could therefore lead to better quality information collected by each member. One objective of this chapter is to illustrate the possibility of each of these effects, and show under what kind of parameter conditions (priors, cost and quality of better information, etc.) they arise.

Larger committees are unambiguously better when the information complementarity effect applies, both because individual members collect superior information and because there are more sources of information. However, if conditions are such that larger committees are prone to the free rider effect, the designer, in choosing committee size, may face a trade-off between the quantity and quality of information. Larger committees will base their decisions on several pieces of low grade information, while smaller ones act on the basis of fewer pieces of high quality information. We show that the free-rider effect can be strong enough to make smaller committees informationally superior in some cases. The model characterizes parametric situations where it is optimal to keep committee size smaller than what is feasible.

Committee design, because of its obvious importance in the process of decision-making, has attracted much attention. One strand of literature extends the "strategic information transmission" model of Crawford and Sobel [14] to examine how a decision-maker's welfare can be influenced by varying the composition of a committee consisting of members with useful information ("expertise"), whose policy preferences differ from that of the decision-maker (or the organization). Gilligan and Krehbiel [26], Austen-Smith [1, 2], Krishna and Morgan [30, 31], Battaglini [5] examine situations where all committee members have the same information but may have different preferences. Wolinsky [52] examines a situation where experts have the same preferences but may receive different signals. Holmstrom [28], Dessein [17], Li and Suen [37], de Garidel-Thoron and Ottaviani [15] concentrated their focus on the effects of delegation in the same setting.

Our focus in this chapter is on the incentive for collecting information by committee members[23]. There are a number of other studies which deal with partly similar issues. But, while we focus on the issue of committee size, most of these studies attempt to see how the incentive for information collection varies with the decision rule in a committee of fixed size. Li, Rosen and Suen [38] and Li [36] examined optimal decision rules in the context of a fixed committee size. Li [35] analyzes the organizational structure that

[23]For an excellent survey of the existing literature on committee design with endogenous information, see Gerling, Grüner, Kiel and Schulte [25].

minimizes information processing costs for a specific task. Dessein [18] also examines a model of organization with conflicting interest to study authoritative coordination vs. consensus in decision making. Sah and Stiglitz [49, 50] investigate similar issues of organizational design, but do not stress incentives to collect information. In their structure, information is exogenous. Blinder and Morgan [7] compare group versus individual decision making in an organization using an experimental study. Haleblian and Finkelstein [27], in an econometric analysis, test how managerial team size affects the organizational decision-making using firm-level data and conclude that in more uncertain environments larger teams tend to do better. In the context of jury trial, Mukhopadhyay [45] identifies the free rider problem in a large jury.

There are two ways in which this analysis departs from the existing literature on committee design that addresses incentives to collect costly information.We consider only *cheap talk* mechanisms. Except for the private cost of collecting information, there is no conflict of interest among committee members or the decision maker in our model, implying that members have no incentive to withhold anything they have learned. We assume the committee is advisory, and the designer cannot commit to ignore any decision relevant information ex-post, which means all the information that has been collected will be efficiently used. The only strategic choice the decision maker faces in our framework is the size of the committee. In contrast, Gerardi and Yariv [24], allow full commitment to any mechanism at the information aggregation stage, including ones which distort the use of available information. Persico [46] analyzes committees that make their decisions through voting, giving the designer the option to choose committee size as well as the voting rule. All these papers assume some commitment power on the part of the designer as to how the information will be used. We think it is natural to assume in many contexts, especially those involving advisory committees, that commitment to distort the use of information once it becomes available is hard to sustain.

We also allow for information of different qualities, and members can choose the precision of the information they gather. Specifically, we assume each member has access to a free informative signal, but can obtain a more precise signal if she pays a private cost. This creates a possible trade-off between quantity and quality of information from the committee designer's point of view. This aspect is absent from the other models[24] of committee design that address similar issues (for example, Persico [46] and Gerardi and Yariv [24]), and has a non-trivial effect on the results. In other papers, larger committees are informationally superior to smaller ones in the weak sense, i.e. it is never the case that a smaller committee generates strictly better information than a larger one. Most papers demonstrate an upper bound to informativeness as committee size goes up. In contrast, we find situations where a smaller committee is informationally superior in the strict sense, which implies that when it comes to committee design, too many cooks may spoil the broth.

One application that fits well with the model proposed in this essay is the example of a hiring committee in an academic department. Hiring com-

[24]Karoutkin and Paroush [29] has examined the quantity versus quality dilemma in an exogenous setting. In their set-up, both the quality and quantity of information are exogenous and as the committee size decreases the quality of information for each member rises automatically. In our model, this occurs as an equilibrium phenomenon (for some parametric configurations).

mittees generally consist of existing faculty members. Once the department agrees on the field of the candidate it is going to hire, the faculty as a whole wants to hire a good candidate. Hence, there is no fundamental conflict of interest among the committee members. The committee members collect information about a potential recruit and then pass on their recommendations to the departmental head or the dean who then takes the final decision based on the recommendations of the committee members.

The committee members get a basic idea about a candidate's qualities by attending the interview and looking at a candidate's CV. A committee member may also collect better information regarding the candidate's qualities by putting in extra effort in reading the research papers authored by the candidate or talking to people who have better idea about the candidate from past experiences. Presumably, the extra effort induces higher personal opportunity cost for the committee member. Once too many members are included in a committee, each individual member's incentive to put in that extra effort would reduce. More specifically, an individual committee member may bank upon others to read the candidate's papers and make the correct recommendations. Since all committee members have common interest, if everybody else puts in the higher effort, this course of action saves the higher opportunity cost for an individual committee member without vastly compromising the efficiency of the decision. But if a large number of committee members follow the same, the efficiency of the decision would be significantly reduced. Our model analyzes a scenario almost similar to above.

The organization of the chapter is as follows. In section 2, we set up the basic model. Section 3 analyzes the model and points out the main results.

Section 4 concludes.

4.2 The Model

An organization has to form a committee to take a decision. There are n individuals from which the committee has to be formed. The decision problem is to choose from among two options— A (alternative) or Q (status-quo). There are two possible states of nature as well - A and Q. The common preference of each individual is such that with complete information the optimal decision is matched exactly with the state. This is captured by the utility function $u(d, \omega)$, where $d \in \{A, Q\}$ represents the ultimate decision and ω is the state-of-the-world. If the decision is correct, each member of the organization gets utility equal to one. The utility from a wrong decision is 0. The preference is captured by the following utility function:

$$u(\omega, d) = \begin{cases} 1 & \text{if } d = \omega \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

The prior on the event $\omega = A$ is π.

There are two signal technologies indexed by $t \in \{h, l\}$. For each technology, the signal can take one of two possible values from the set $\Theta_t = \{a_t, q_t\}$. A signal of type t has an accuracy p_t i.e. for a $t-$type signal, $\Pr[a_t|A] = \Pr[q_t|Q] = p_t \in (\frac{1}{2}, 1)$ for $t = h, l$. We also assume that the signals are conditionally independent.

We assume the l-technology is costless. This means that even if an individual does not invest in acquiring information, she still gets a noisy signal by default. By investing in an effort amounting to cost c, each individual can collect a signal of better quality. The better signal has a precision level p_h with $p_h > p_l$.

For the sake of simplicity, we will restrict ourselves to two possible values of n - 1 or 2. Increasing n to values greater than 2 does not add much insight towards understanding the main points, but complicates the analysis. We model the situation in the following way. The committee designer first chooses the committee size n (1 or 2). The member(s) of the committee then decide simultaneously whether to collect the better signal by incurring the private cost or just depend on the lower quality signal. The members report their signals to the committee designer who takes the decision after utilizing all the available information. One assumption we make here is that the committee designer cannot make any credible commitment regarding not to use any available information at the decision stage. This implies that once the signals are collected, they are utilized optimally. In other words, information is fully aggregated prior to the decision. Everybody shares a common utility function; hence there is no incentive to withhold information under this mechanism after the collection stage.

The assumption regarding full information aggregation merits some discussion. We do not take up the optimal mechanism design problem. Notice that if we assume the committee designer can credibly commit to any mechanism, then the smaller committee outcome can always be mimicked with a larger committee in which the designer commits to ignore the messages of some players. But there always remains a question regarding implementability of such mechanisms. Given common interest between the decision maker and the committee members it is hard to imagine that the decision maker can credibly commit to an ex post decision rule that promises to use less information than what is available to him. After all, we do not often see the heads of the departments or the deans of the faculties handing out explicit contracts to hiring committee members which say that not all recommendations would be used in decision making. We, therefore, restrict ourselves to a particular symmetric mechanism namely the ex post efficient mechanism[25]. However, we recognize, as has been pointed out by Gerardi and Yariv [24], for large enough n, the ex post efficient mechanism may not be the optimal mechanism. But that does not contradict our result. Even if we choose the optimal symmetric mechanism instead of the full information aggregation mechanism as the decision rule, a smaller committee can perform strictly better than a larger committee under certain parameter values. The intuitions behind the results remain similar.

The problem we take up here is not one of finding the optimal committee size with an infinite number of potential members available. Obviously, the optimal committee size in this case is infinite. Since the default signal is somewhat informative (even if slightly), in a very large committee the probability of making the correct decision is almost equal to one. One way to interpret this model is to consider it as a constrained problem. Given that only a finite number of people are available for inclusion in the committee, we discuss how committees should be formed.

For comparison across committees, we use the common expected value generated by the information collected by the committee exclusive of in-

[25]The ex post efficient mechanism is one where the committee designer uses all the available information at the decision making stage

formation costs. This we think is the natural welfare function when the preferences are common. If the decision affects a very large group (or an organization), the individual costs are small or negligible relative to the welfare of the people affected. Henceforth, we denote this common value generated by a n-member committee by $V^n(.)$.

Before proceeding further, we need to impose some restrictions on the parameters of the model. We assume that $p_h - p_l > c$. We need this restriction to make the better signals attractive at least in some cases. For the rest of the chapter, we restrict ourselves to values of $\pi \geq \frac{1}{2}$. Similar results can be obtained for $\pi \leq \frac{1}{2}$ because of the symmetric nature of the model.

Some discussions regarding the choice of this model are worthwhile here. The standard models of committee design with endogenous information (for example, Persico [46] and Gerardi and Yariv [24]) consider a two level choice of information quality. In these models, the committee members have access to either a high quality signal or no information at all. This dampens the incentive for free riding since the alternative has no informative value. That is why these models cannot generate a strict dominance result for smaller committee size. Our model, though apparently also has a two level choice of signal quality, is actually the first approximation of a model with richer quality choice. Similar qualitative results can be obtained if we introduce a cost (which must be appropriately low) for the lower quality signal. This suggests that the result we obtain here will survive in a more general model with multi-level quality choice.

4.3 Committee Design

Let θ be the vector of signals collected by all players before the decision is made. Since each individual reports their signals truthfully, θ is completely known prior to decision-making. Then the common posterior on the state being A, defined by Bayes' Rule, is as follows:

$$\gamma(\theta) = \Pr[\omega = A|\theta] = \frac{\pi \Pr(\theta|\omega = A)}{\pi \Pr(\theta|\omega = A) + (1-\pi)\Pr(\theta|\omega = Q)}. \tag{4.2}$$

Clearly, under the no commitment mechanism, the optimal decision $d^*(\theta)$ is as follows:

$$d^*(\theta) = A \Longleftrightarrow \gamma(\theta) \geq \frac{1}{2}. \tag{4.3}$$

We can write the ex-post common utility of this decision as

$$U(\gamma(\theta)) = \begin{cases} \gamma(\theta), \text{ if } \gamma(\theta) \geq \frac{1}{2} \\ 1-\gamma(\theta), \text{ if } \gamma(\theta) < \frac{1}{2} \end{cases}. \tag{4.4}$$

Let $P(\theta)$ be the probability of realization of a signal vector θ. Possible realizations of θ of course depend on the chosen signal technologies. Let $T \in \{h, l\}^2$ be the vector of chosen signal technologies and Θ_T be the set of all possible signal realizations under T. Then ex-ante common expected value from T can be written as

$$V(T) = \sum_{\theta \in \Theta_T} P(\theta) U(\gamma(\theta)). \tag{4.5}$$

We are now in a position to discuss the equilibrium outcomes for different committee sizes. We take up that task in the next two subsections.

4.3.1 One member committee

We first consider the single person decision problem. Without loss of generality, we concentrate on the range of priors $\pi \in [\frac{1}{2}, 1]$. Because of the symmetry of the structure, exactly similar results can be obtained for $\pi \in [0, \frac{1}{2})$. For the rest of the chapter, we assume $\pi \geq \frac{1}{2}$. A t-type signal, $t \in \{h, l\}$ is informationally decision relevant if and only if $\pi \in [\frac{1}{2}, p_t)$. In case the prior falls in this range, a signal q_t pushes the posterior below $\frac{1}{2}$ and the decision is contingent on the realization of the signal. Hence the expected utility from collecting a signal of type t without accounting for cost is given by

$$V^1(t) = P(q_t)\left(1 - \gamma\left(q_t\right)\right) + P\left(a_t\right)\gamma\left(a_t\right)$$

for $\pi \in [\frac{1}{2}, p_t)$ and $t = l, h$. For $\pi \geq p_t$, the optimal decision is A independent of the signal realization. Therefore, $V^1(t) = \pi$ for $\pi \geq p_t$. Hence, the marginal benefit from collecting a h-type signal in the single person decision problem is

$$b^1(\pi) = \begin{cases} p_h - p_l & \text{if } \frac{1}{2} \leq \pi < p_l \\ p_h - \pi & \text{if } p_l \leq \pi < p_h \\ 0 & \text{otherwise} \end{cases} . \tag{4.6}$$

Notice that $b^1(\pi)$ is a continuous function of π. The individual collects h if and only if $b^1(\pi) \geq c$. This condition induces an interval of priors over which the h-type signal is collected by the individual. We summarize the finding in the following lemma.

Lemma 4.1. *For any $c < p_h - p_l$, in a one member committee, a h-type signal will be collected if and only if $\pi \in [\frac{1}{2}, p_h - c]$.*

Proof. Since $b^1(\pi)$ is strictly decreasing in π for all $\pi \in [p_l, p_h]$, and reaches 0 at p_h, for any $c > 0$, $b^1(\pi) \geq c$ induces a unique interval for any $c < p_h - p_l = \max b^1(\pi)$. ∎

The social value function from this individual decision (ignoring the cost) is then given by[26]

$$V^1(\pi) = \begin{cases} p_h & \text{if } \pi \leq p_h - c \\ \pi & \text{otherwise} \end{cases}. \tag{4.7}$$

Since $p_h - c > p_l$, the l-type signal has no decision relevance in the range where the h-type signal is not being collected and hence cannot affect the social value.

One observation may be worth mentioning here. Notice that as p_l falls, $p_h - p_l$ rises. As we mentioned earlier, $p_h - p_l$ indicates the upper bound on the range of cost parameter for which an individual can be induced to collect a high quality signal. This shows that the presence of a free informative signal may sometimes lower the incentive for providing effort to find a better signal. In other words, free information may be expensive from a social point of view.

4.3.2 Two member committee

Next we move on to our analysis of a two member committee. Let $t_i \in \{h, l\}$ denote player i's choice of signal technology and $V^2(t_i, t_j)$ represent the payoff to players when players i and j choose t_i and t_j respectively. For different

[26]We are slightly abusing our notation here. We write the value function assuming that optimal decision d^* is taken for all ranges of prior, given the incentive constraint for the individual. We do this throughout the chapter for the sake of comparison across the committees.

parameter values, three types of pure strategy equilibria may exist in a two member committee: (i) (h, h), (ii) (h, l) or (l, h) and (iii) (l, l). We next characterize the necessary and sufficient conditions for existence of these pure strategy equilibria in terms of the value function.

1. $V^2(h, h) - V^2(h, l) \geq c$ *is necessary and sufficient for* (h, h) *to be an equilibrium.*

2. $V^2(h, h) - V^2(h, l) < c$ *and* $V^2(h, l) - V^2(l, l) \geq c$ *are necessary and sufficient for* (h, l) *and* (l, h) *to be equilibria.*

3. $V^2(h, l) - V^2(l, l) < c$ *is necessary and sufficient for* (l, l) *to be an equilibrium.*

These conditions follow directly from the definition of Bayesian Nash Equilibrium which is the equilibrium concept we use throughout this essay. Since we are in an environment with common preferences where information is a public good, players' ex-ante utilities (net of cost) depend only on the types of signal technologies chosen by them, but not on the exact combination of choices. Hence, the value functions mentioned above are sufficient statistics for characterizing the equilibria in this environment. An immediate corollary of the above conditions is that at least one pure strategy equilibrium exists for all parameter values.

The necessary and sufficient conditions for existence of different pure strategy equilibria also show that neither (h, h) nor (l, l) can coexist with (h, l) or (l, h) as equilibria. Moreover, for parameter values such that (h, l) is an equilibrium, (l, h) is an equilibrium as well. Our next lemma proves that when (l,l) or (h,h) is the unique pure strategy equilibrium, no mixed strategy equilibrium can exist.

Lemma 4.2. *For parameter values such that (l,l) or (h,h) is the unique pure strategy equilibrium, then it is the unique equilibrium of the two person game.*

Proof. First consider that (l,l) is the unique pure strategy equilibrium. Then, $V^2(h,h) - V^2(h,l) < c$ and $V^2(h,l) - V^2(l,l) < c$. Now consider any mixed strategy σ for player j where σ is the probability of playing h. Generically, for σ to be part of a mixed strategy equilibrium, we must have the following

$$\sigma V^2(h,h) + (1-\sigma) V^2(h,l) - c = \sigma V^2(l,h) + (1-\sigma) V^2(l,l)$$

or, equivalently

$$\sigma \left[V^2(h,h) - V^2(l,h)\right] + (1-\sigma) \left[V^2(h,l) - V^2(l,l)\right] = c.$$

From the two strict inequalities described above, and the fact that $V^2(h,l) = V^2(l,h)$, it follows that, for any $\sigma \in [0,1]$, the lhs of the above is strictly less than c. Hence, we cannot have a mixed strategy equilibrium in this case.

The proof for the case when (h,h) is the unique pure strategy equilibrium is similar. ∎

For any signal vector θ, we define $\Pi(\theta)$ to be such that for all $\pi \geq \Pi(\theta)$, $\gamma(\theta) \geq \frac{1}{2}$ where $\gamma(\theta)$ is the common posterior on the state being A given θ. We first consider the case when both players collect h. For any π[27], $\gamma(a_h, a_h)$,

[27] Remember that we are restricting ourselves to values of π greater than or equal to $\frac{1}{2}$.

$\gamma(a_h, q_h)$ and $\gamma(q_h, a_h)$ are greater than or equal to $\frac{1}{2}$ and the optimal decision corresponding to these signal realizations is A. But $\gamma(q_h, q_h) < \frac{1}{2}$ if and only if $\pi < \Pi(q_h, q_h)$ and in that case the decision is Q under full information utilization. Hence, in the case where (h, h) are the signal technologies chosen by the players, the signals have decision relevance if and only if $\pi < \Pi(q_h, q_h)$. The expected utility to each player when both players choose h is the following:

$$
\begin{aligned}
& V^2(h,h) \\
= & \begin{cases} \begin{array}{c} P(a_h, a_h)\gamma(a_h, a_h) + P(a_h, q_h)\gamma(a_h, q_h) \\ +P(q_h, a_h)\gamma(q_h, a_h) + P(q_h, q_h)\left[1 - \gamma(q_h, q_h)\right] \end{array} & \text{if } \pi < \Pi(q_h, q_h) \\ \pi & \text{otherwise} \end{cases} \\
= & \begin{cases} p_h^2 + 2p_h(1-p_h)\pi & \text{if } \pi < \Pi(q_h, q_h) \\ \pi & \text{otherwise} \end{cases}. \qquad (4.8)
\end{aligned}
$$

Similarly, when both players choose l, the common expected utility is

$$
V^2(l,l) = \begin{cases} p_l^2 + 2p_l(1-p_l)\pi & \text{if } \pi < \Pi(q_l, q_l) \\ \pi & \text{otherwise} \end{cases}. \qquad (4.9)
$$

When the chosen signal technologies are (h, l) or (l, h), the derivation of the common expected utility is a little more involved. The l-type signal has practical relevance in decision-making if and only if the prior falls in a range such that a $l-$signal in favour of A coupled with the prior overwhelms a $h-$signal in favour of Q to clinch the decision in favour of A. For this to happen, π must belong to the interval $[\Pi(a_l, q_h), \Pi(q_l, q_h))$. For $\pi <$

neither signal matters. Hence the common expected utility can be written as

$$V^2(h,l) = \begin{cases} p_h & \text{if } \pi < \Pi(a_l, q_h) \\ p_h p_l + [p_h(1-p_l) + p_l(1-p_h)]\,\pi & \text{if } \Pi(a_l, q_h) \le \pi < \Pi(q_l, q_h) \\ \pi & \text{otherwise} \end{cases} . \tag{4.10}$$

Let $b^2(\pi; t)$ denote the marginal benefit from collecting an $h-$type signal when the other player is collecting a $t-$type signal. Obviously, we can now write

$$b^2(\pi; h) = V^2(h,h) - V^2(h,l) \tag{4.11}$$

and

$$b^2(\pi; l) = V^2(h,l) - V^2(l,l). \tag{4.12}$$

Since a better signal technology cannot reduce the ex-ante value of information, $b^2(\pi, t) \ge 0$ for all π. Both $b^2(\pi; h)$ and $b^2(\pi; l)$ are continuous functions of π. The following two lemmas characterize the parameter zones for which incentive constraints for collecting the $h-$type signal conditional on the other player's strategy are satisfied.

Lemma 4.3. *For any $c < p_h - p_l$, there exists a unique $\bar{\pi}_l(c) < \Pi(q_l, q_h)$ such that $b^2(\pi; l) \ge c$ if and only if $\pi \le \bar{\pi}_l(c)$.*

Proof. See Appendix A.2. ∎

Lemma 4.4. *Fix p_h, p_l. There exist $c_0 > 0$ such that for all $c < c_0$, there exist unique cutoffs $\pi_0^h(c)$ and $\pi_1^h(c)$ such that $b^2(\pi; h) \ge c$ if and only if $\pi \in [\pi_0^h(c), \pi_1^h(c)]$ where $\pi_0^h(c) < \Pi(a_l, q_h)$ and $\Pi(a_l, q_h) < \pi_1^h(c) < \Pi(q_h, q_h)$.*

Proof. See Appendix A.2. ∎

As a corollary to Lemma 4.4, we can see that a necessary condition for a (h, h) equilibrium is $c < c_0$. As the cost of the $h-$type signal increases, the free-rider effect becomes more intense and it is more difficult to satisfy the incentive constraints for both individuals collecting the $h-$type signals.

4.3.3 Comparison between committees

We are now in a position to make comparisons across committees with respect to effort levels and effect on general welfare. We state our main results in two propositions. The first proposition identifies parameter zones where the domination of the complementarity effect over the free-rider effect (and vice versa) can be clearly seen. The second proposition provides conditions for the strict domination of a smaller committee over a larger committee.

We define a set $\Phi = \{\pi \in [p_l, p_h] : b^1(\pi) = b^2(\pi; l)\}$. The following lemma characterizes the set Φ.

Lemma 4.5. *Φ is non-empty, closed and bounded.*

Proof. Define $g(\pi) = b^1(\pi) - b^2(\pi; l)$. $g(\pi)$ is continuous since both $b^1(\pi)$ and $b^2(\pi; l)$ are continuous. $b^2(\pi; l)$ is strictly decreasing over the domain $\left[\frac{1}{2}, \Pi(q_l, q_h)\right]$ and $b^2\left(\frac{1}{2}; l\right) = p_h - p_l$. But, $b^1(p_l) = p_h - p_l$ and $b^1(p_h) = 0$. Hence $g(p_l) > 0$. and $g(p_h) < 0$. An application of Intermediate Value Theorem then ensures that Φ is non-empty. The boundedness obviously follows from definition of Φ. To see that Φ is closed, assume the contrary. Then there exists a sequence $\{\pi_n\}$ such that $g(\pi_n) = 0$ for all n, but $\lim_{n\to\infty} g(\pi_n) \neq 0$. But this contradicts continuity of $g(.)$. ∎

Since $\Phi \subset \Re$, the last Lemma ensures that both $\max_\pi \Phi$ and $\min_\pi \Phi$ exist and are unique. Let us denote these two by $\pi^l_{\max}$ and $\pi^l_{\min}$ respectively. Now define $c_1 = b^1\left(\pi^l_{\max}\right)$ and $c_2 = b^1\left(\pi^l_{\min}\right)$. Since $b^1(.)$ is strictly decreasing in this range, $c_1 \leq c_2$. An inspection of $b^1(.)$ and $b^2(.;l)$ reveals that we essentially confront two possible scenarios. If $\Pi(a_l, q_h) < \Pi(q_l, q_l)$, then $\pi^l_{\max} = \pi^l_{\min} \in (\Pi(a_l, q_h), \Pi(q_l, q_l))$. On the other hand, if $\Pi(a_l, q_h) \geq \Pi(q_l, q_l)$, then $\pi^l_{\min} = \Pi(q_l, q_l)$ and $\pi^l_{\max} = \Pi(a_l, q_h)$.

Finally, for any c, define $\bar{\pi} = \max\{\pi^1_h(c), \bar{\pi}_l(c)\}$, where $\pi^1_h(c)$ and $\bar{\pi}_l(c)$ are as defined in Lemma 4.3 and Lemma 4.4 respectively. Let T^*_n denote the vector of equilibrium signal technologies chosen by a committee of size n. Now, we can state our first proposition.

Proposition 4.1. *1. $c < c_1$ and $\pi \in (p_h - c, \bar{\pi}]$ are sufficient for $T^*_1 = l$, and existence of a $T^*_2 \in \{(h,h), (h,l), (l,h)\}$.*

2. $c > c_2$ and $\pi \in (\bar{\pi}_l(c), p_h - c]$ are necessary and sufficient for $T^*_1 = h$ and $T^*_2 = (l,l)$.

Proof. See Appendix A.2. ∎

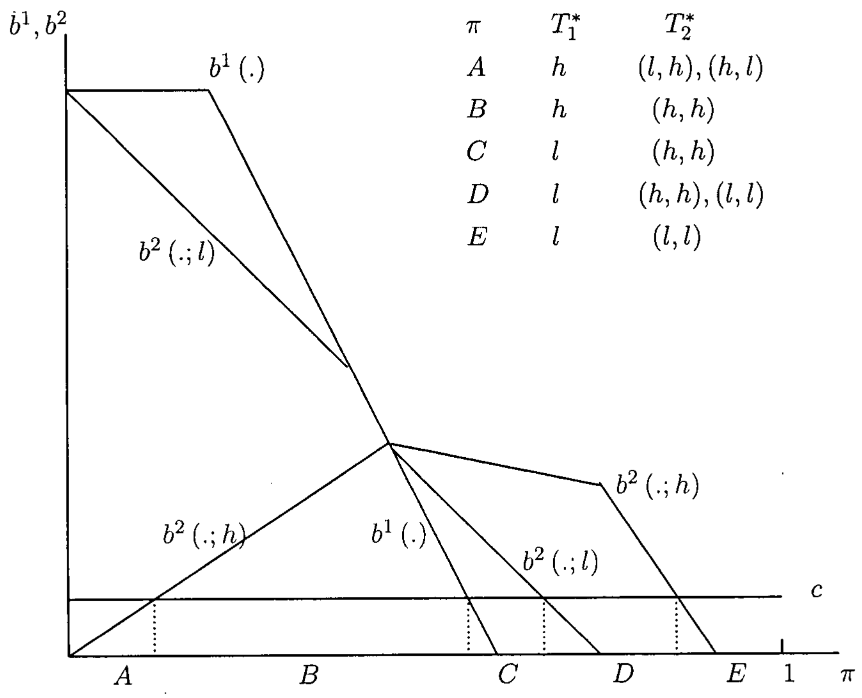
Proposition 4.1 describes the effect of committee size on incentives. It identifies conditions under which the larger committee is subject to either the free rider effect or the information complementarity effect, and proves neither of these are empty. The first part provides a sufficient condition for information complementarity effect to come into play. Under these conditions, the equilibrium signal quality is l in a one member committee, but in a two member committee, there always exists an equilibrium with at least one h-type signal. The second part of the proposition provides necessary and sufficient conditions for the free-rider effect to apply. It identifies parameter zones where a one member committee collects a h-type signal in equilibrium, while in a two member committee the only equilibrium is (l, l).

The result is quite intuitive. The incentive to free-ride is more intense when the cost of quality information is high. On the other hand, signals become complementary towards more extreme priors, since the cumulative signals must be sufficiently strong to have any relevance to the decision. Hence, for the complementarity effect, we need a sufficiently low value of c along with relatively large prior values. Exactly the opposite is true for the free-rider effect. We need a sufficiently high c along with not so extreme priors for this to happen.

Some discussion regarding the set of equilibria for different parameter values is worthwhile here. In the following two figures, we illustrate these in the case of a particular set of parameter values, namely when $\Pi(q_l, q_l) < \Pi(a_l, q_h)$. The other cases can be dealt with similarly.

We illustrate the above proposition with the help of a numerical example. For parameter values $p_h = 0.8$ and $p_l = 0.6$, the values of c_1 and c_2 in the above proposition can be easily computed to be 0.0727 and 0.1077 respectively. Figures 4.1 and 4.2 are drawn with parameter values $p_h = 0.8$ and $p_l = 0.6$ to show the pure strategy equilibria for different values of prior.

In Figure 4.1, we take $c = 0.02 < c_1$. The panel at the top right hand corner of Figure 4.1 identifies the pure strategy equilibria when the prior falls in different zones. For π close to $\frac{1}{2}$ ($\pi \in A = [0.5000, 0.5625)$), the bigger committee admits (h, l) and (l, h) equilibria. For extreme values of π ($\pi \in E = (0.9118, 1.0000]$), neither type of committee puts in an ef-

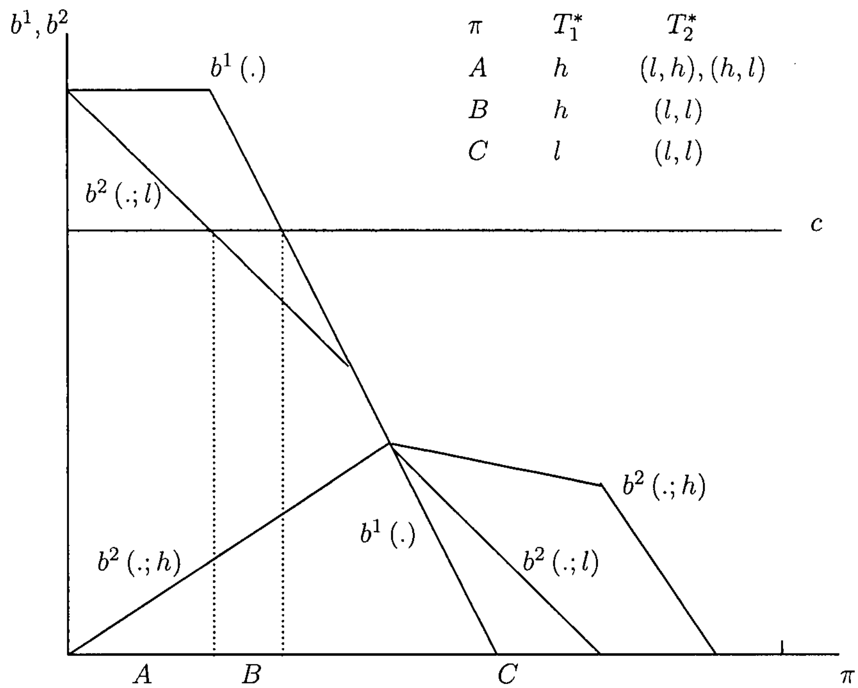


Note: The above figure is drawn with parameter values $p_l = 0.6$, $p_h = 0.8$ and $c = 0.02$.

Figure 4.1: Pure Strategy Equilibria for Different Prior Values: $c < c_1$.

fort to collect a $h-$type signal. For other values of π, the bigger committee admits a (h, h) equilibrium. When π falls in $C = (0.7800, 0.8214]$ or $D = (0.8214, 0.9118]$, one (h, h) equilibrium exists in a two member committee while a single member committee does not put any effort at all into gathering the better signal[28]. For values of $\pi \in B = [0.5625, 0.7800]$, the bigger committee collects two $h-$type signals while the single member com-

[28]This is just for the purpose of illustration. We can have other parameter configurations satisfying the conditions mentioned in Proposition 4.1 such that values of prior exist where a one member committee collects l in equilibrium while a two member committee admits (h, l) or (l, h) as equilibria.



Note: The above figure is drawn with parameter values $p_l = 0.6$, $p_h = 0.8$ and $c = 0.15$.

Figure 4.2: Pure Strategy Equilibria for Different Prior Values: $c > c_2$.

mittee collects only one in equilibrium. B, C and D are the zones where a clear domination of the complementarity effect over the free-rider effect can be seen as identified in Proposition 1.

Figure 4.2 illustrates the set of equilibria for $c = 0.15 > c_2$. Given that $c > c_2$, for priors very close to $\frac{1}{2}$ ($\pi \in A = [0.5000, 0.6042]$), $T_1^* = h$ and $T_2^* = (h,l)$ or (l,h). Both these equilibria generate the same social value[29] as the single member committee for this particular situation. Then, we have a

[29] Of course, there exist mixed strategy equilibria for the bigger committee in this range, which may do strictly better than the smaller committee outcome.

range of priors ($\pi \in B = (0.6042, 0.6500]$) where $T_1^* = h$ and $T_2^* = (l, l)$. This is the parameter zone we identified in part 2 of Proposition 4.1. Towards the extreme prior values belonging to the interval $C = (0.6500, 1.0000]$, $T_1^* = l$ and $T_2^* = (l, l)$.

Notice that in both cases illustrated above, for values of prior close enough to $\frac{1}{2}$, $T_1^* = h$, and $T_2^* = (h, l)$ or (l, h), that is the smaller and the bigger committees are informationally equivalent. In other words, both committees generate the same level of decision relevant information and hence social value. This is what Persico [46] and Gerardi and Yariv [24] identified.

Their results show that given the parameters, there exists an upper bound on the social value that can be generated by increasing committee size. In their models, in the absence of an exogenous cost of designing a bigger committee, a smaller committee cannot do strictly better. In our next proposition, we provide conditions for the strict dominance result of the smaller committee.

In part 2 of Proposition 4.1 we have identified the conditions under which a one member committee collects h in equilibrium, while the unique equilibrium in the two member committee is (l, l). This is necessary for welfare dominance of the smaller committee, but may not be sufficient. In our next proposition, we show that for all values of p_h and p_l, we can find parameter zones where $T_1^* = h$ and $T_2^* = (l, l)$ become sufficient for strict welfare dominance of the smaller committee.

Proposition 4.2. *For all p_l,$p_h \in (\frac{1}{2}, 1)$ such that $p_h > p_l$, there exist π and c such that the smaller committee dominates the larger committee welfarewise.*

Proof. See Appendix A.2. ∎

We illustrate again with numerical examples. We show in the proof of Proposition 4.2 that if $p_h \geq \frac{p_l^2}{p_l^2+(1-p_l)^2}$, $T_1^* = h$ and $T_2^* = (l,l)$ are sufficient for strict welfare dominance of the smaller committee. For parameter values $p_h = 0.8$ and $p_l = 0.6$, $p_h \geq \frac{p_l^2}{p_l^2+(1-p_l)^2}$. For these values of p_h and p_l, $c_2 = 0.1077$ and $\pi_{\min}^l = 0.6923$ can be easily computed. Now, at $\pi = 0.6923$, $V^2(l,l) = 0.6923$. For any $c > c_2$, the values of π for which $T_1^* = h$ and $T_2^* = (l,l)$ are strictly less than 0.6923. Since $V^2(l,l)$ is increasing in π, hence $V^2(l,l) < V^1(h) = 0.8$ in the relevant zone.

Next we consider parameter values $p_h = 0.8$ and $p_l = 0.7$. Now, p_h is strictly less than $\frac{p_l^2}{p_l^2+(1-p_l)^2}$. We choose $c = 0.05$ such that (h,h) is not an equilibrium. In the range of prior given by $(0.6190, 0.7500)$, $T_1^* = h$ and $T_2^* = (l,l)$. To show that this no longer suffices for strict welfare dominance of the smaller committee, we choose $\pi = 0.74$ where $V^2(l,l) = 0.8008$, but $V^1(h) = 0.8$. But for lower values of π, in particular for values of π in the interval $(0.6190, 0.7381)$, $V^1(h) > V^2(l,l)$ still holds.

Notice that so far we have chosen the common value of each individual for welfare comparisons. Suppose instead we choose a utilitarian welfare function that sums up each individual's net welfare to obtain social welfare. As before a necessary condition for welfare dominance of the smaller committee is that the smaller committee chooses to collect a h-type signal while the larger committee collects only l-type signals. As shown in Proposition 4.1 this can happen only if $\pi \in (\bar{\pi}_l(c), p_h - c]$ and $c > c_2$. Notice that social welfare if information is collected by a smaller committee is now $2V^1(h) - c$, while in case of the larger committee social welfare is $2V^2(l,l)$. Suppose we again take $p_h = 0.8$ and $p_l = 0.6$. As we have already shown, for these values of

p_h and p_l, $c_2 = 0.1077$ and $\pi^l_{\min} = 0.6923$. Suppose we choose $c = 0.15$. Then the highest value of π for which $T_1^* = h$ and $T_2^* = (l,l)$ is 0.65. At $\pi = 0.65, 2V^2(l,l) = 1.344$. Since $V^2(l,l)$ is increasing in π, $2V^2(l,l) < 1.344$ in the relevant zone, while $2V^1(h) - c = 1.45$. Hence, for these parameter values the smaller committee dominates the bigger committee welfarewise even for utilitarian welfare function. This example illustrates the robustness of our result with respect to the specification of the welfare function.

For high enough c, the incentive to free ride dominates the positive incentive towards good quality signals arising out of complementarity between the signals in a larger committee, at least for some values of the prior. In a smaller committee, the lower free riding incentive may induce the individuals to collect a larger number of high quality signals, making the smaller committee better from a social point of view. This indicates the quality - quantity trade-off that the committee designer faces in an environment with multiple information qualities.

Notice that the last proposition is a strict violation of the Condorcet Jury Theorem of the first kind. Even under the assumption of full information aggregation, a smaller committee strictly dominates a bigger committee. In a two level signal quality choice model with the alternative to the informative signal being a completely uninformative one, this result cannot be obtained. It can be easily seen in this model by verifying that as $p_l \to \frac{1}{2}, b^1(\pi) \to b^2(\pi; l)$ for all values of π.

4.4 Concluding Remarks

This essay doesn't attempt to give a general characterization of the solution to optimal committee design problems. Previous analyzes have provided a weak dominance result for smaller committees in that there is an upper bound on the value of information generated by increasing the committee size. We, on the other hand, argue that not allowing committee members choice over the quality of information they collect misses an important aspect of the problem. We show that in a slightly richer model which allows for this, there are situations when a smaller committee can do strictly better, even under the assumption of perfect information aggregation. We expect this result to be generic and robust to model specifications.

Bibliography

[1] Austen-Smith, David (1990): "Information Transmission in Debate." *American Journal of Political Science* **34:** 124-52.

[2] Austen-Smith, David (1993): "Interested Experts and Policy Advice: Multiple Referrals under Open Rule." *Games and Economic Behavior* **5:** 3-43.

[3] Austen-Smith, David and Jeffrey S. Banks (1996): "Information Aggregation, Rationality and the Condorcet Jury Theorem." *American Political Science Review* **90**: 34–45.

[4] Barro, R. J., and D. Gordon (1983): "Rules, discretion, and reputation in a model of monetary policy." *Journal of Monetary Economics* **12:** 101 21.

[5] Battaglini, Marco (2002): "Multiple Referrals and Multidimensional Cheap Talk." *Econometrica* **70:** 1379-401.

[6] Berend, Daniel and Jacob Paroush (1988): "When is Condorcet's Jury Theorem Valid?" *Social Choice and Welfare* **15:** 481-88.

[7] Blinder, Alan S. and John Morgan (2000): "Are Two Heads Better Than One?: An Experimental Analysis of Group vs. Individual Decision-making." *NBER Working Paper* **No. 7909.**

[8] Cai, H. (2001): "Optimal Committee Design With Heterogeneous Preferences." Mimeo, UCLA.

[9] Callander, Steven (2002): "Vote Timing and Information Aggregation." Mimeo, Northwestern University.

[10] Chakraborty, Archishman and Parikshit Ghosh (2002): "Efficient Equilibria and Information Aggregation in Common Interest Voting Games." Mimeo, University of British Columbia.

[11] Chamley, Christopher and Douglas Gale (1994): "Information Revelation and Strategic Delay in a Model of Investment." *Econometrica* **62:** 1065-85.

[12] Condorcet, Marquis de (1785): "Essay on the Application of Mathematics to the Theory of Decision-Making." Reprinted in Keith Michael Baker (ed.): *Condorcet: Selected Writings*, Indianapolis: Bobbs-Marrill Co.

[13] Coughlan, Peter (2000): "In Defence of Unanimous Jury Verdicts: Mistrials, Communication and Strategic Voting." *American Political Science Review* **94**: 375–93.

[14] Crawford, Vincent P. and Joel Sobel (1982): "Strategic Information Transmission." *Econometrica* **50:** 1431-51.

[15] de Garidel-Thoron, Thomas and Marco Ottaviani (2000): "The Economics of Advice." Mimeo, University College of London.

[16] Dekel, Eddie and Michele Piccione (2000): "Sequential Voting Procedures in Symmetric Binary Elections." *Journal of Political Economy* **108:** 34-55.

[17] Dessein, Wouter (2002a): "Authority and Communication in Organizations." *Review of Economic Studies* **69:** 811-38.

[18] Dessein, Wouter (2002b): "Coordination Through Authority versus Consensus." Manuscript, University of Chicago.

[19] Doraszelski, Ulrich, Dino Gerardi and Francesco Squintani (2003): "Communication and Voting with Double-sided Information." *Contributions to Theoretical Economics* **3**: No. 1, Article 6.

[20] Duggan, John and Cesar Martinelli (1999): "A Bayesian Model of Voting in Juries." *Games and Economic Behavior* **37**: 259–94.

[21] Feddersen, Timothy and Wolfgang Pesendorfer (1997): "Voting Behaviour and Information Aggregation in Elections with Private Information." *Econometrica* **65**: 1029–58.

[22] Feddersen, Timothy and Wolfgang Pesendorfer (1998): "Convicting the Innocent: the Inferiority of Unanimous Jury Verdicts." *American Political Science Review* **92**: 23–35.

[23] Gerardi, Dino and Leeat Yariv (2003): "Putting your Ballot Where Your Mouth Is - Analysis of Collective Choice with Communication." UCLA Working Paper, No. 827.

[24] Gerardi, Dino and Leeat Yariv (2004): "Mechanism Design with Information Acquisition." Manuscript, Yale University.

[25] Gerling, Kirsten, Hans Peter Grüner, Alexandra Kiel and Elisabeth Schulte (2003): "Information Acquisition and Decision Making in Committees: A Survey." European Central Bank Working Paper, No. 256

[26] Gilligan, Thomas W. and Keith Krehbiel (1989): "Asymmetric Information and Legislative Rules with a Heterogeneous Committee." *American Journal of Political Science* **33:** 459-90.

[27] Haleblian, Jerayr and Sydney Finkelstein (1993): "Top Management Team Size, CEO Dominance, and Firm Performance: The Moderating Roles of Environmental Turbulence and Discretion." *Academy of Management Journal* **36:** 844-63.

[28] Holmström, B. (1984): "On the Theory of Delegation." In *Bayesian Models in Economic Theory,* ed. by Marcel Boyer and Richard Khilstrom. Elsevier Science Publishers B.V.

[29] Karotkin, Drora and Jacob Paroush (2003): "Optimum committee size: Quality-versus-quantity dilemma." *Social Choice and Welfare* **20:** 429-41.

[30] Krishna, Vijay and John Morgan (2001a): "A Model of Expertise." *Qarterly Journal of Economics* **116:** 747-75.

[31] Krishna, Vijay and John Morgan (2001b): "Assymmetric Information and Legislative Rules: Some Amendments." *American Political Science Review* **95**: 435-52.

[32] Kydland, F. E., and E. Prescott (1997): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy* **85**: 47391.

[33] Ladha, Krishna (1992): "The Condorcet Jury Theorem, Free Speech and Correlated Votes." *American Journal of Political Science* **36:** 617-34.

[34] Ladha, Krishna, Gary Miller and Joe Oppenheimer (2000): "Information Aggregation by Majority Rules: Theory and Experiments." Mimeo, University of Maryland.

[35] Li, Hao (1999): "Hierarchies and Information-Processing Organizations." *Review of Economic Design* **4:** 101-26.

[36] Li, Hao (2001): "A Theory of Conservatism." *Journal of Political Economy***109:** 617-36.

[37] Li, Hao and Wing Suen (2001): "Delegating Decisions to Experts." *Journal of Political conomy,* forthcoming.

[38] Li, Hao, Sherwin Rosen and Wing Suen (2001): "Conflicts and Common Interests in Committees." *American Economic Review* **91:** 1478-97.

[39] Martinelli, Cesar (2002): "Convergence Results for Unanimous Voting." *Journal of Economic Theory* **105**: 278-97.

[40] McLean, I. and F. Hewitt (1994): Condorcet: Foundations of social choice and political theory. Elgar, Brook.eld, VT

[41] McLennan, Andrew (1998): "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents." *American Political Science Review* **92**: 413–18.

[42] Meirowitz, Adam (2001): "Informative Voting and Condorcet Theorems with a Continuum of Types." Mimeo, Stanford Graduate School of Business.

[43] Miller, Nicholas (1986): "Information, Electorates and Democracy: Some Extensions and Interpretations of the Condorcet Jury Theorem." In *Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine, Conference on Political Economy,* ed. by Bernard Grofman and Guillarmo Owen. Greenwich, CT: JAI Press.

[44] Morton, Rebecca B. and Kenneth C. Williams (2000): *Learning by Voting: Sequential Choices in Presidential Primaries and Other Elections.* Ann Arbor, Michigan: University of Michigan Press.

[45] Mukhopadhyaya, Kalyan (1998): "Jury Size and the Free Rider Problem." Mimeo, Indiana University, Bloomington.

[46] Persico, Nicola (2004): "Committee Design with Endogenous Information." *Review of Economic Studies,* forthcoming.

[47] Radner, Roy and Joseph Stiglitz (1984): "A Nonconcavity in the Value of Information." In *Bayesian Models in Economic Theory,* ed. by Michael Boyer and Richard Khilstrom. Elsevier Science Publishers B.V.

[48] Rogoff, K. (1985): "The Optimal Degree of Commmitment to a Monetary Target." *Quarterly Journal of Economics* **100:** 116989.

[49] Sah, Raaj Kumar and Joseph Stiglitz (1986): "The Architecture of Economic Systems: Hierarchies and Polyarchies." *American Economic Review* **76:** 716-27.

[50] Sah, Raaj Kumar and Joseph Stiglitz (1988): "Committees, Hierarchies and Polyarchies." *Economic Journal* **98:** 451-70.

[51] Wit, Jorgen (1998): "Rational Choice and Condorcet Jury Theorem." *Games and Economic Behavior* **22:** 364-76.

[52] Wolinsky, Asher (2002): "Eliciting Information from Multiple Experts." *Games and Economic Behavior* **41**: 141-60.

[53] Young, Peyton (1986): "Optimal Voting Rules." *Journal of Economic Perspectives* (Winter) **9**: 51-64.

[54] Young, Peyton (1988): "Condorcet's Theory of Voting." *American Political Science Review* **82**: 1231-44.

A. Appendices

A.1 Appendix to Chapter 2

A.1.1 Proof of Lemma 2.3

Proof. We first express $V(m+2;q_D) - V(m;q_D)$ explicitly in terms of the parameters as the following:

$$\begin{aligned} V(m+2;q_D) - V(m;q_D) = & - \sum_{x=0}^{R_s(m)} \binom{m+2}{x} p^x (1-p)^{m+2-x} \pi (1-q_D) \\ & - \sum_{x=R_s(m)+1}^{m+2} \binom{m+2}{x} p^{m+2-x} (1-p)^x (1-\pi) q_D \\ & + \sum_{x=0}^{R_s(m)-1} \binom{m}{x} p^x (1-p)^{m-x} \pi (1-q_D) \\ & + \sum_{x=R_s(m)}^{m} \binom{m}{x} p^{m-x} (1-p)^x (1-\pi) q_D . \quad \text{(A.1)} \end{aligned}$$

To prove this lemma we first show that for a generic $1 \leq R \leq m$,

$$\sum_{x=0}^{R-1} \binom{m}{x} p^x (1-p)^{m-x} - \sum_{x=0}^{R} \binom{m+2}{x} p^x (1-p)^{m+2-x}$$
$$= \binom{m+1}{R} (1-p)^{m+1-R} p^R \left\{ \frac{R}{m+1} - (1-p) \right\} \quad \text{(A.2)}$$

and

$$\sum_{x=R}^{m} \binom{m}{x} p^{m-x} (1-p)^x - \sum_{x=R+1}^{m+2} \binom{m+2}{x} p^{m+2-x} (1-p)^x$$
$$= \binom{m+1}{R} p^{m+1-R} (1-p)^R \left\{ p - \frac{R}{m+1} \right\}. \quad \text{(A.3)}$$

We use the method of induction to prove the above results. For notational convenience we write

$$\phi(R) = \sum_{x=0}^{R-1} \binom{m}{x} p^x (1-p)^{m-x} - \sum_{x=0}^{R} \binom{m+2}{x} p^x (1-p)^{m+2-x}$$

and

$$\psi(R) = \sum_{x=R}^{m} \binom{m}{x} p^{m-x} (1-p)^x - \sum_{x=R+1}^{m+2} \binom{m+2}{x} p^{m+2-x} (1-p)^x.$$

For $R = 1$, the first equality holds. That both LHS and RHS of equation [A.2] are equal to $p(1-p)^m - (m+1)p(1-p)^{m+1}$ can be verified after some algebraic manipulation. Now suppose that the equality holds for some arbitrary R. We show that then the equality holds for $R+1$. We need to show that

$$\phi(R+1) = \binom{m+1}{R+1}(1-p)^{m-R} p^{R+1} \left\{ \frac{R+1}{m+1} - (1-p) \right\} \tag{A.4}$$

But,

$$\begin{aligned}
\phi(R+1) &= \phi(R) + \binom{m}{R} p^R (1-p)^{m-R} - \binom{m+2}{R+1} p^{R+1} (1-p)^{m-R+1} \\
&= \binom{m+1}{R} (1-p)^{m+1-R} p^R \left\{ \frac{R}{m+1} - (1-p) \right\} \\
&+ \binom{m}{R} p^R (1-p)^{m-R} - \binom{m+2}{R+1} p^{R+1} (1-p)^{m-R+1}
\end{aligned}$$

by induction hypothesis. After cancelling $(1-p)^{m-R} p^R$ from equation [A.4] and some simplification we are left to prove the following:

$$\begin{aligned}
\binom{m}{R-1}(1-p) &- \binom{m+1}{R}(1-p)^2 + \binom{m}{R} - \binom{m+2}{R+1} p(1-p) \\
&= \binom{m}{R} p - \binom{m+1}{R+1} p(1-p)
\end{aligned}$$

which can be shown to be true using the fact that $\binom{n+1}{x} - \binom{n}{x} = \binom{n}{x-1}$.

Next we consider the second equality. That equation [A.3] holds for $R = m$ can be seen from verifying that both LHS and RHS of [A.3] are equal to $p^2(1-p)^m - mp(1-p)^{m+1}$. Suppose equation [A.3] holds for some arbitrary $R+1$. We will show that then it will hold for R. The induction hypothesis tells us that $\psi(R+1) = \binom{m+1}{R+1} p^{m-R} (1-p)^{R+1} \left\{p - \frac{R+1}{m+1}\right\}$. Now,

$$\begin{aligned}\psi(R) &= \psi(R+1) + \binom{m}{R} p^{m-R} (1-p)^{R} - \binom{m+2}{R+1} p^{m-R+1} (1-p)^{R+1} \\ &= \binom{m+1}{R+1} p^{m-R} (1-p)^{R+1} \left\{p - \frac{R+1}{m+1}\right\} \\ &+ \binom{m}{R} p^{m-R} (1-p)^{R} - \binom{m+2}{R+1} p^{m-R+1} (1-p)^{R+1}.\end{aligned}$$

We need to show that

$$\psi(R) = \binom{m+1}{R} p^{m+1-R} (1-p)^{R} \left\{p - \frac{R}{m+1}\right\}. \tag{A.5}$$

After cancelling $p^{m-R}(1-p)^R$ from both sides of equation [A.5] and some simplification we are left to prove

$$\begin{aligned}\binom{m+1}{R+1} p(1-p) &- \binom{m}{R}(1-p) + \binom{m}{R} - \binom{m+2}{R+1} p(1-p) \\ &= \binom{m+1}{R} p^2 - \binom{m}{R-1} p\end{aligned}$$

which can be shown to be true again using the relation $\binom{n+1}{x} - \binom{n}{x} = \binom{n}{x-1}$.

Since equations [A.2] and [A.3] hold for a generic R, we can now write $V(m+2; q_D) - V(m; q_D)$ in the following manner:

$$\begin{aligned}& V(m+2; q_D) - V(m; q_D) \\ =\ & \pi(1-q_D)\phi(R_s(m)) + (1-\pi) q_D \psi(R_s(m)) \\ =\ & \binom{m+1}{R_s(m)} [(1-p)^{m+1-R_s(m)} p^{R_s(m)} \left\{\frac{R_s(m)}{m+1} - (1-p)\right\} \pi(1-q_D) \\ +\ & p^{m+1-R_s(m)} (1-p)^{R_s(m)} \left\{p - \frac{R_s(m)}{m+1}\right\} (1-\pi) q_D]\end{aligned} \tag{A.6}$$

Suppose that $V(m;q_D) - V(m-2;q_D) > 0$. We first consider the case when $\frac{R_s(m)}{m+1} \leq \frac{1}{2} < p$. If $\frac{R_s(m)}{m+1} \geq 1-p$, then $V(m+2;q_D) - V(m;q_D) > 0$. Suppose $\frac{R_s(m)}{m+1} < 1-p$. Then $V(m+2;q_D) - V(m;q_D) > 0$ if and only if

$$\begin{aligned}
\frac{p^{m+1-R_s(m)}(1-p)^{R_s(m)}(1-\pi)q_D}{(1-p)^{m+1-R_s(m)}p^{R_s(m)}\pi(1-q_D)} &> \frac{1-p-\frac{R_s(m)}{m+1}}{p-\frac{R_s(m)}{m+1}} \\
\Leftrightarrow \frac{p^{m-R_s(m)}(1-p)^{R_s(m)-1}(1-\pi)q_D}{(1-p)^{m-R_s(m)}p^{R_s(m)-1}\pi(1-q_D)} &> \frac{1-p-\frac{R_s(m)}{m+1}}{p-\frac{R_s(m)}{m+1}}.
\end{aligned}$$

But,

$$\frac{R_s(m)}{m+1} \leq \frac{1}{2} \Rightarrow \frac{R_s(m)-1}{m-1} \leq \frac{R_s(m)}{m+1} \leq \frac{1}{2}.$$

Hence, using $R_s(m-2) = R_s(m) - 1$ from Lemma 2.2 in equation [A.6] and writing $m-2$ in place of m we find

$$\begin{aligned}
& V(m;q_D) - V(m-2;q_D) \\
= & \binom{m-1}{R_s(m)-1}[(1-p)^{m-R_s(m)}p^{R_s(m)-1}\left\{\frac{R_s(m)-1}{m-1} - (1-p)\right\}\pi(1-q_D) \\
+ & p^{m-R_s(m)}(1-p)^{R_s(m)-1}\left\{p - \frac{R_s(m)-1}{m-1}\right\}(1-\pi)q_D] \\
> & 0 \\
\Rightarrow & \frac{p^{m-R_s(m)}(1-p)^{R_s(m)-1}(1-\pi)q_D}{(1-p)^{m-R_s(m)}p^{R_s(m)-1}\pi(1-q_D)} > \frac{1-p-\frac{R_s(m)-1}{m-1}}{p-\frac{R_s(m)-1}{m-1}}.
\end{aligned}$$

Since $\frac{1-p-\frac{R_s(m)-1}{m-1}}{p-\frac{R_s(m)-1}{m-1}} \geq \frac{1-p-\frac{R_s(m)}{m+1}}{p-\frac{R_s(m)}{m+1}}$ when $\frac{R_s(m)-1}{m-1} \leq \frac{R_s(m)}{m+1}$,

$$V(m;q_D) - V(m-2;q_D) > 0 \Rightarrow V(m+2;q_D) - V(m;q_D) > 0.$$

Now consider the other possibility that $\frac{R_s(m)}{m+1} > \frac{1}{2} > 1-p$. If $\frac{R_s(m)}{m+1} \leq p$, then $V(m+2;q_D) - V(m;q_D) > 0$. Suppose $\frac{R_s(m)}{m+1} > p$. Then $V(m+2;q_D) - V(m;q_D) > 0$ if and only if

$$\frac{p^{m+1-R_s(m)}(1-p)^{R_s(m)}(1-\pi)q_D}{(1-p)^{m+1-R_s(m)}p^{R_s(m)}\pi(1-q_D)} < \frac{\frac{R_s(m)}{m+1}-(1-p)}{\frac{R_s(m)}{m+1}-p}$$

$$\Leftrightarrow \frac{p^{m-R_s(m)}(1-p)^{R_s(m)-1}(1-\pi)q_D}{(1-p)^{m-R_s(m)}p^{R_s(m)-1}\pi(1-q_D)} < \frac{\frac{R_s(m)}{m+1}-(1-p)}{\frac{R_s(m)}{m+1}-p}.$$

But, $\frac{R_s(m)}{m+1} > \frac{1}{2} \Rightarrow \frac{R_s(m)-1}{m-1} > \frac{R_s(m)}{m+1} > \frac{1}{2}$. Hence,

$$V(m;q_D) - V(m-2;q_D) > 0$$
$$\Rightarrow \frac{p^{m-R_s(m)}(1-p)^{R_s(m)-1}(1-\pi)q_D}{(1-p)^{m-R_s(m)}p^{R_s(m)-1}\pi(1-q_D)} < \frac{\frac{R_s(m)-1}{m-1}-(1-p)}{\frac{R_s(m)-1}{m-1}-p}.$$

Since $\frac{\frac{R_s(m)}{m+1}-(1-p)}{\frac{R_s(m)}{m+1}-p} > \frac{\frac{R_s(m)-1}{m-1}-(1-p)}{\frac{R_s(m)-1}{m-1}-p}$ when $\frac{R_s(m)-1}{m-1} > \frac{R_s(m)}{m+1}$,

$$V(m;q_D) - V(m-2;q_D) > 0 \Rightarrow V(m+2;q_D) - V(m;q_D) > 0$$

in this case as well. This completes the proof of the lemma. ∎

A.1.2 Proof of Lemma 2.4

Proof. We have to deal with two cases separately:

Case 1: $\rho_J > 1$; The jury puts relatively higher weight on acquitting the guilty.

Case 2: $\rho_J < 1$. The jury puts relatively higher weight on convicting the innocent.

Case 1: $\rho_J > 1$.

The uninformative decision of the jury in this case is C. Hence, $V(0; q_D) = -(1-\pi)q_D$.

Now suppose n_0 is the smallest positive integer such that a jury of size n_0 votes informatively. Then $R_s(n_0) = R_s(n_0+1) = 1$. For any $n < n_0$, the jury votes uninformatively for C. Hence, $R_s(n_0) = 1$. Now $R_s(n_0+1) = 1$ or 2 by Lemma 2.2. For any $n_0 > 1$, if $R_s(n_0+1) = 2$, then again by Lemma 2.2, $R_s(n_0-1) = 1$ contradicting the hypothesis that n_0 is the smallest integer such that the jury votes informatively. For $n_0 = 1$, $R_s(2) = 1$ follows directly from the fact that $\rho_J > 1$.

We will argue that $V(n_0; q_D) > V(0; q_D) \Rightarrow V(n_0+2; q_D) > V(n_0; q_D)$ and $V(n_0+1; q_D) > V(0; q_D) \Rightarrow V(n_0+3; q_D) > V(n_0+1; q_D)$. Then an application of Lemma 2.3 will be sufficient for the proof in this case.

Suppose $V(n_0; q_D) > V(0; q_D)$. Then, using $R_s(n_0) = 1$, we can show that $(1-\pi)\, q_D p^{n_0} > \pi(1-q_D)(1-p)^{n_0}$ follows. As shown in the proof of Lemma 2.3, we can write $V(n_0+2; q_D) - V(n_0; q_D)$ as

$$
\begin{aligned}
& V(n_0+2; q_D) - V(n_0; q_D) \\
= & \binom{n_0+1}{1} \left[\begin{array}{c} \pi(1-q_D)(1-p)^{n_0} p \left(\frac{1}{n_0+1} - (1-p)\right) \\ +(1-\pi)\, q_D p^{n_0} (1-p) \left(p - \frac{1}{n_0+1}\right) \end{array} \right] \\
> & \binom{n_0+1}{1} \pi(1-q_D)(1-p)^{n_0} \left[(1-p)\left(p - \frac{1}{n_0+1}\right) + p\left(\frac{1}{n_0+1} - (1-p)\right)\right]. \\
> & 0
\end{aligned}
$$

where we made use of the fact $(1-\pi)\, q_D p^{n_0} > \pi(1-q_D)(1-p)^{n_0}$ and $\frac{1}{n_0+1} \leq \frac{1}{2} < p$.

The subcase for $V(n_0+1;q_D) > V(0;q_D)$ can be treated similarly.

Case 2: $\rho_J < 1$.

The uninformative decision of the jury in this case is A. Hence, $V(0;q_D) = -\pi(1-q_D)$.

Again, as in Case 1, suppose n_0 is the smallest positive integer such that a jury of size n_0 votes informatively. Then $R_s(n_0) = n_0$ and $R_s(n_0+1) = n_0+1$. For any $n < n_0$, the jury votes uninformatively for A. Hence, $R_s(n_0) = n_0$. Now $R_s(n_0+1) = n_0+1$ or n_0 by Lemma 2.2. For any $n_0 > 1$, if $R_s(n_0+1) = n_0$, then again by Lemma 2.2, $R_s(n_0-1) = n_0-1$ contradicting the hypothesis that n_0 is the smallest integer such that the jury votes informatively. For $n_0 = 1$, $R_s(2) = 2$ follows directly from the fact that $\rho_J < 1$.

Now we can proceed similarly as in Case 1 to show that $V(n_0;q_D) > V(0;q_D) \Rightarrow V(n_0+2;q_D) > V(n_0;q_D)$ and $V(n_0+1;q_D) > V(0;q_D) \Rightarrow V(n_0+1;q_D) > V(n_0+3;q_D)$. ∎

A.2 Appendix to Chapter 4

A.2.1 Proof of Lemma 4.3

Proof. In writing $b^2(\pi;l)$ explicitly in terms of the parameters, we have to consider two separate cases.

Case I: $\Pi(a_l,q_h) \geq \Pi(q_l,q_l)$

Using equations [4.10] and [4.9], we can write

$$b^2(\pi;l) = \begin{cases} p_h - p_l^2 - 2p_l(1-p_l)\pi & \text{if } \pi < \Pi(q_l,q_l) \\ p_h - \pi & \text{if } \Pi(q_l,q_l) \leq \pi < \Pi(a_l,q_h) \\ p_h p_l - [p_h p_l + (1-p_h)(1-p_l)]\pi & \text{if } \Pi(a_l,q_h) \leq \pi < \Pi(q_l,q_h) \\ 0 & \text{otherwise.} \end{cases}.$$

Hence, in this case, $b^2(\pi;l)$ is continuous and strictly decreasing in π for all $\frac{1}{2} \leq \pi \leq \Pi(q_l,q_h)$.

Case II: $\Pi(a_l,q_h) < \Pi(q_l,q_l)$

Again using equations [4.10] and [4.9], $b^2(\pi;l)$ can be expressed as

$$b^2(\pi;l) = \begin{cases} p_h - p_l^2 - 2p_l(1-p_l)\pi & \text{if } \pi < \Pi(a_l,q_h) \\ (p_h - p_l)[p_l - (2p_l - 1)\pi] & \text{if } \Pi(a_l,q_h) \leq \pi < \Pi(q_l,q_l) \\ p_h p_l - [p_h p_l + (1-p_h)(1-p_l)]\pi & \text{if } \Pi(q_l,q_l) \leq \pi < \Pi(q_l,q_h) \\ 0 & \text{otherwise.} \end{cases}.$$

Here also, $b^2(\pi;l)$ is continuous and strictly decreasing in π for all $\frac{1}{2} \leq \pi \leq \Pi(q_l,q_h)$.

Hence, irrespective of the parameter configurations, $b^2(\pi;l)$ is monotonically decreasing in π. Moreover, $b^2(\pi;l) = p_h - p_l > 0$ at $\pi = \frac{1}{2}$ and $b^2(\pi;l) = 0$ at $\pi = \Pi(q_l,q_h)$. Therefore, any $0 < c < p_h - p_l$ will induce a unique cutoff $\bar{\pi}_l(c) \in \left(\frac{1}{2}, \Pi(q_l,q_h)\right)$ such that $b^2(\pi;l) \geq c$ if and only if $\pi \leq \bar{\pi}_l(c)$. ∎

A.2.2 Proof of Lemma 4.4

Proof. We first express $b^2(\pi; h)$ in terms of the parameters of the model using equations [4.8] and [4.10]:

$$b^2(\pi; h) = \begin{cases} p_h^2 + 2p_h(1-p_h)\pi - p_h & \text{if } \pi < \Pi(a_l, q_h) \\ (p_h - p_l)[p_h(1-\pi) + (1-p_h)\pi] & \text{if } \Pi(a_l, q_h) \le \pi < \Pi(q_l, q_h) \\ p_h^2(1-\pi) + (1-p_h)^2\pi & \text{if } \Pi(q_l, q_h) \le \pi < \Pi(q_h, q_h) \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $b^2(\pi; h)$ is throughout continuous in π and has a unique maxima at $\Pi(a_l, q_h)$. Hence, any $c < \max_\pi b^2(\pi; h) = \frac{(p_h - p_l)p_h(1-p_h)}{p_h(1-p_l)+(1-p_h)p_l} = c_0$, induces an interval $[\pi_0^h(c), \pi_1^h(c)]$ such that for $\pi \in [\pi_0^h(c), \pi_1^h(c)]$, $b^2(\pi; h) \ge c$. Since $b^2(\frac{1}{2}; h) = b^2(\Pi(q_h, q_h); h) = 0$ and $b^2(\Pi(a_l, q_h); h) = c_0$, $c < c_0$ implies $\pi_0^h(c) \in (\frac{1}{2}, \Pi(a_l, q_h))$ and $\pi_1^h(c) \in (\Pi(a_l, q_h), \Pi(q_h, q_h))$. ∎

A.2.3 Proof of Proposition 4.1

Proof. First Part:

For any c, $T_1^* = l$ for $\pi > p_h - c$ can be seen from Lemma 4.1. For $c < c_1$, $\bar{\pi}_l(c) > p_h - c$. Hence from Lemma 4.4, (l, l) cannot be an equilibrium for $\pi \in (p_h - c, \bar{\pi}_l(c)]$. Since we already argued that a pure strategy equilibrium always exists, T_2^* must be (h, l) or (l, h) or (h, h) in this zone. Now consider the case where $\pi_h^1(c) > \bar{\pi}_l(c)$. For $\pi \in (\bar{\pi}_l(c), \pi_h^1(c)]$, $b^2(\pi; l) < c$, but $b^2(\pi, h) \ge c$. Hence, in this zone we have a (l, l) equilibrium, but we also have (h, h) as an equilibrium.

Second Part:

Sufficiency

Suppose $c > c_2$. Then $\pi \leq p_h - c$ is sufficient for $T_1^* = h$ by Lemma 4.1. For $c > c_2$, $\bar{\pi}_l(c) < p_h - c$. Hence, for $\pi \in (\bar{\pi}_l(c), p_h - c]$, (h, l) or (l, h) cannot be an equilibrium by Lemma 4.4. We now need to show that (h, h) cannot be an equilibrium in this zone. We have to consider two separate cases.

Case I: $\Pi(a_l, q_h) \geq \Pi(q_l, q_l)$

In this case, $\pi^l_{\min} = \Pi(q_l, q_l)$. Hence,

$$\begin{aligned} c_2 &= b^1(\Pi(q_l, q_l)) \\ &= p_h - \Pi(q_l, q_l) \\ &\geq p_h - \Pi(a_l, q_h) \\ &= p_h - \frac{p_h(1-p_l)}{p_h(1-p_l) + p_l(1-p_h)} \\ &= \frac{p_h(1-p_h)(2p_l - 1)}{p_h(1-p_l) + p_l(1-p_h)}. \end{aligned}$$

Since $b^2(\pi; h)$ is strictly increasing in π for all $\pi \in \left[\frac{1}{2}, \Pi(a_l, q_h)\right]$, $b^2(\Pi(q_l, q_l); h) < c_2$ along with $\Pi(q_l, q_l) \leq \Pi(a_l, q_h)$ implies that for any $\pi < \Pi(q_l, q_l)$ and $c > c_2$, $b^2(\pi; h) < c$. The required condition can be easily verified as follows:

$$\begin{aligned} b^2(\Pi(q_l, q_l); h) &= p_h(1-p_h)(2\Pi(q_l, q_l) - 1) \\ &= \frac{p_h(1-p_h)(2p_l - 1)}{p_l^2 + (1-p_l)^2} \\ &< \frac{p_h(1-p_h)(2p_l - 1)}{p_h(1-p_l) + p_l(1-p_h)} \\ &\leq c_2. \end{aligned}$$

Since for any $c > c_2$, $p_h - c < \Pi(q_l, q_l)$, hence for any $\pi \leq p_h - c$, (h, h) cannot be an equilibrium. We know that for all parameter configurations a pure strategy equilibrium exists. Hence, (l, l) is the only pure strategy equilibrium which is the unique equilibrium by virtue of Lemma 4.2.

Case II: $\Pi(a_l, q_h) < \Pi(q_l, q_l)$

Here $c_2 = b^2\left(\pi^l_{\min}\right)$ where $\pi^l_{\min} \in (\Pi(a_l, q_h), \Pi(q_l, q_l))$. Hence, for any $\pi \in \left[\Pi(a_l, q_h), \pi^l_{\min}\right]$,

$$b^2(\pi; l) = (p_h - p_l)\left[p_l(1-\pi) + (1-p_l)\pi\right]$$

and

$$b^2(\pi; h) = (p_h - p_l)\left[p_h(1-\pi) + (1-p_h)\pi\right].$$

Since $\pi > \frac{1}{2}$ in this range, $b^2(\pi; l) > b^2(\pi; h)$ for all $\pi \in \left[\Pi(a_l, q_h), \pi^l_{\min}\right]$. Moreover, since $b^2(\pi; l)$ is strictly decreasing for all π and $b^2(\pi; h)$ reaches its maximum at $\Pi(a_l, q_h)$, $b^2(\pi; l) > b^2(\pi; h)$ for all $\pi \in \left[\frac{1}{2}, \pi^l_{\min}\right]$. Hence, $\pi > \bar{\pi}_l(c)$ implies $b^2(\pi; h) < c$. Therefore, (h, h) cannot be an equilibrium outcome.

Since we know that for all parameter configurations a pure strategy equilibrium exists, $T_2^* = (l, l)$ is the only pure strategy equilibrium for $\pi \in (\bar{\pi}_l(c), p_h - c]$, which is the unique equilibrium by virtue of Lemma 4.2.

Necessity

For $c > c_2$, $\pi > p_h - c$ implies $T_1^* = l$ and $\pi \leq \bar{\pi}_l(c)$ implies T_2^* includes at least one h by Lemma 4.4. For $c \leq c_2, \bar{\pi}_l(c) \geq p_h - c$ for all c and hence $T_1^* = h$ implies T_2^* includes at least one h, again by Lemma 4.4. ∎

A.2.4 Proof of Proposition 4.2

Proof. It is evident that $T_1^* = h$ and $T_2^* = (l,l)$ are necessary for the smaller committee to dominate the larger committee. For any $c > c_2$, the existence of such situation is guaranteed by Proposition 4.1. In our relevant zone, $V^1(h) > V^2(l,l)$ if and only if $\pi < \frac{p_h - p_l^2}{2p_l(1-p_l)}$. First consider that $p_h \geq \frac{p_l^2}{p_l^2+(1-p_l)^2}$. Then it follows that $\frac{p_h-p_l^2}{2p_l(1-p_l)} \geq \frac{p_l^2}{p_l^2+(1-p_l)^2}$. Since $\pi^l_{\min} \leq \frac{p_l^2}{p_l^2+(1-p_l)^2}$, therefore $\pi^l_{\min} \leq \frac{p_h-p_l^2}{2p_l(1-p_l)}$. Now, for any $c > c_2$, $p_h - c < \pi^l_{\min} \leq \frac{p_h-p_l^2}{2p_l(1-p_l)}$. Hence, for parameter values such that $p_h \geq \frac{p_l^2}{p_l^2+(1-p_l)^2}$, $T_1^* = h$ and $T_2^* = (l,l)$ are sufficient for $V^1(h) > V^2(l,l)$.

Now suppose $p_h < \frac{p_l^2}{p_l^2+(1-p_l)^2}$. This implies that $\frac{p_h-p_l^2}{2p_l(1-p_l)} < \frac{p_l^2}{p_l^2+(1-p_l)^2}$. If $\pi^l_{\min} \leq \frac{p_h-p_l^2}{2p_l(1-p_l)}$, then by the same argument as in the preceding paragraph, $T_1^* = h$ and $T_2^* = (l,l)$ are sufficient for $V^1(h) > V^2(l,l)$. Now suppose $\pi^l_{\min} > \frac{p_h-p_l^2}{2p_l(1-p_l)}$. Notice that $\frac{p_h-p_l^2}{2p_l(1-p_l)} > \frac{p_l-p_l^2}{2p_l(1-p_l)} = \frac{1}{2}$. We have defined $\bar{\pi}_l(c)$ such that $b^2(\bar{\pi}_l(c); l) = c$. Since $b^2(\pi; l)$ is continuous, strictly decreasing in π, $\lim_{\pi \to \frac{1}{2}} b^2(\pi; l) = p_h - p_l$ and $b^2(\pi^l_{\min}; l) = c_2$, there exists $c \in (c_2, p_h - p_l)$ such that $\bar{\pi}_l(c) \in \left(\frac{1}{2}, \frac{p_h-p_l^2}{2p_l(1-p_l)}\right)$. Now for any $\pi \in \left(\bar{\pi}_l(c), \min\left\{\frac{p_h-p_l^2}{2p_l(1-p_l)}, p_h - c\right\}\right)$, $T_1^* = h$ and $T_2^* = (l,l)$ and $V^1(h) > V^2(l,l)$. This completes the proof of the proposition. ∎